Shifting Fortunes and Enduring Poverty in Madagascar

RECENT FINDINGS

WORLD BANK GROUP

Shifting Fortunes and Enduring Poverty in Madagascar

RECENT FINDINGS

This report was produced by the Poverty Global Practice, Africa Region, under the supervision of Pablo Fajnzylber. The Task Team Leaders were Theresa Osborne (Senior Economist GPV01) and Nadia Belhaj Hassine Belghith (Senior Economist GPV01).

Regional Vice President: Makhtar Diop
Country Director: Mark Lundell
Country Manager: Coralie Gevers

Contents

Photos

Cover: Gudkov Andrey/Shutterstock; *p. vi:* Danm12/Shutterstock; *p. 3:* Dennis van de Water/Shutterstock; *p. 7:* Damien Ryszawy/Shutterstock; *p. 8:* Agota Kadar/Shutterstock; *p. 12:* Dudarev Mikhail/Shutterstock; *p. 38:* Stock photo/Shutterstock; *p. 58:* Anton Ivanov/Shutterstock; *p. 61:* Dudarev Mikhail/Shutterstock; *p. 65:* Elisabeth/Flickr © All rights reserved; *p.69:* Dennis van de Water/Shutterstock; *p. 84:* Milosk50/Shutterstock; *p. 98:* Dennis van de Water/Shutterstock; *p. 100:* Artush; *p. 106:* Olivier S/Shutterstock; *p. 123:* Damien Ryszawy/Shutterstock; *p. 138:* Pierre Jean Durieu/Shutterstock; *p. 140:* Danm12/Shutterstock; *p. 158:* Anton Ivanov/Shutterstock; *back cover:* Dietmar Temps/Shutterstock

Acronyms

AIF	allocative inefficiency factor
CART	classification and regression tree
ENEMPSI	Enquête nationale sur l'emploi et le secteur informel (National employment and informal sector survey)
ENSOMD	Enquête Nationale sur les Objectifs Millenaire du Développement (National Survey on the Millennium Development Goals)
EPM	Enquête Périodique auprès des Ménages (Periodic Household Survey)
FAOSTAT	Food and Agriculture Statistics
HH	Household
INSTAT	Institut nationale de la statistique. (National Institute of Statistics of Madagascar)
LFS	labor force survey
MRP	marginal revenue product
MSE	mean-squared error
NFE	nonfarm enterprise
NPK	Nitrogen, Phosphorus, Potassium (fertilizer)
OLS	ordinary least squares
OOME	owner-operated microenterprises
PPP	purchasing power parity
RF	random forest
RT	regression tree
SSA	Sub-Saharan Africa
TLU	tropical livestock unit
WDI	World Development Indicators

INTRODUCTION

Poverty and Employment in Madagascar 2001–2012: A Synthesis of Recent Findings

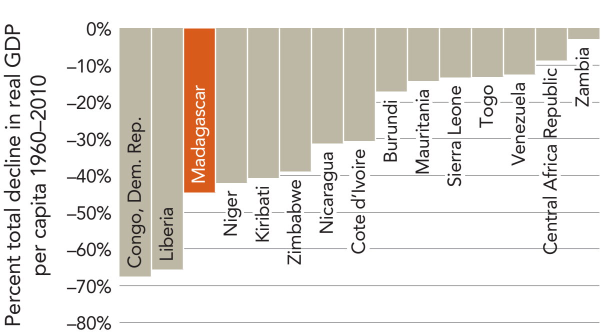
Theresa Osborne
June 2016
AFR Poverty Practice, World Bank

Madagascar remains among the poorest countries in the world, and has shown little improvement in indicators of the well-being of its population over recent years. Despite its unique biodiversity and abundant mineral, water,[1] and labor resources, it ranks among the relatively few countries in the world with real per capita gross domestic product (GDP) in 2010 lower than it was in 1960. Only the Democratic Republic of the Congo and Liberia, two countries which have undergone periods of civil war, have experienced a greater decline (figure I.1). As a result, Madagascar rates as the poorest country in Sub-Saharan Africa (SSA) (and the world) where internationally comparable data are available (figure I.2 maps poverty rates in SSA). This poverty is associated with low and declining labor productivity. By 2012, Madagascar's GDP per employed worker had fallen to the lowest in the world with the exception of the Democratic Republic of the Congo (figure I.3).[2]

Madagascar's economy faces an array of challenges in reducing poverty, including an unfavorable investment climate, severe infrastructure deficits, and political instability (World Bank 2015). In addition, from 2001 to 2012, Madagascar experienced two political crises (in 2002 and 2009); the loss of valuable trade preferences, with the 2005 end of the multifiber agreement, and the 2009 revocation of African Growth and Opportunity Act (AGOA) preferences;[3] and a number of severe droughts, cyclones, and other natural shocks.

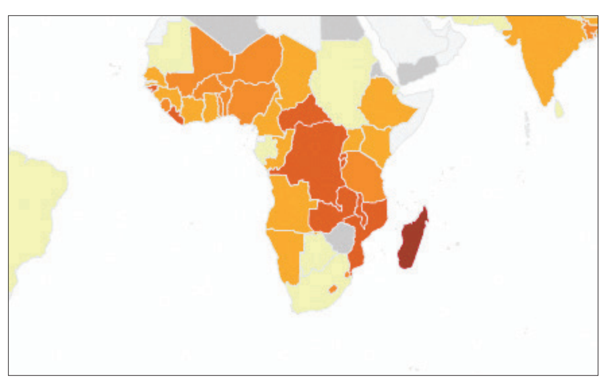
This report synthesizes the insights obtained from a series of five papers on poverty, inequality, labor markets, and returns to agricultural and nonfarm enterprises in Madagascar over the period 2001–12. These papers draw on a combination of empirical techniques, household living standards data, and firm-level data to elucidate key dynamics and structural issues driving poverty and welfare (in all cases measured as per capita

FIGURE I.1: The Countries of the World (with Data Available) Showing Lower Real GDP per Capita in 2010 than 1970



Source: World Development Indicators (WDI).

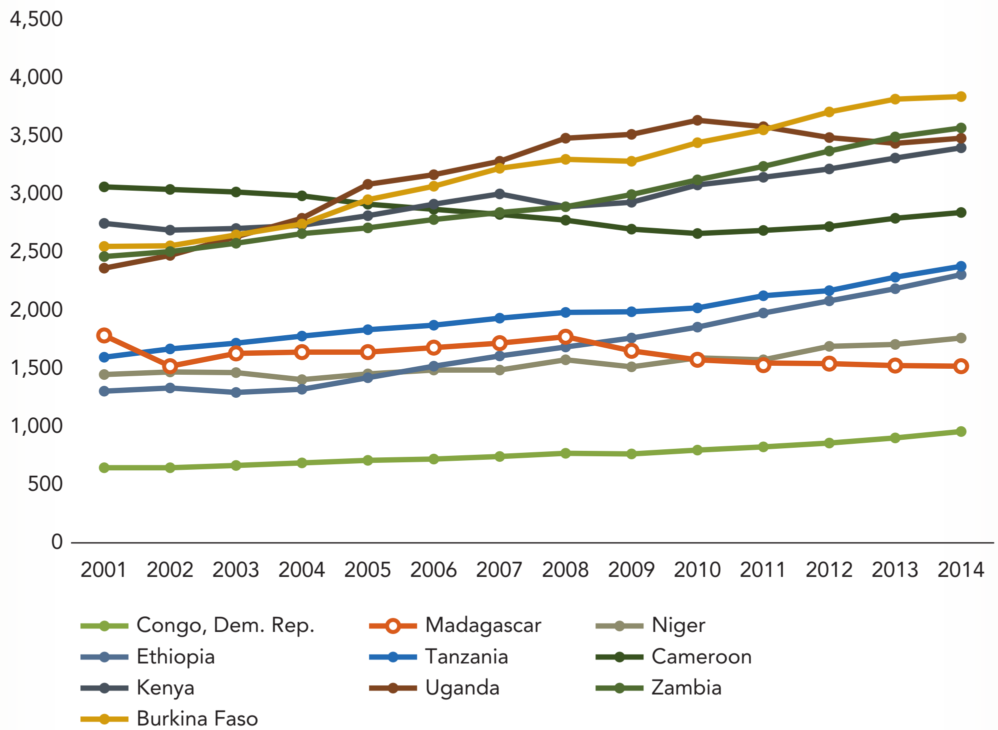
FIGURE I.2: Headcount Poverty Rates, SSA



Yellow-orange = higher rates of extreme poverty.

Source: PovcalNet.

FIGURE I.3: GDP per Employed Person (Constant $US, 1990 PPP)



Source: WDI 2016.
Note: PPP = purchasing power parity.

consumption) over this dozen-year period. First, "Madagascar Poverty and Inequality Update: Recent Trends in Welfare, Employment, and Vulnerability" (Belghith, Randriankolona, and Osborne 2016) updates recent poverty and inequality trends since the World Bank (2014) publication *Face of Poverty in Madagascar: Poverty, Gender and Inequality Assessment*. It also documents trends in key outcomes in both agriculture and labor markets. Second, "Isolation, Crisis, and Vulnerability: A Decomposition Analysis of Inequality and Deepening Poverty in Madagascar (2005–2010)" (Thiebaud, Osborne, and Belghith 2016) uses re-centered influence function estimation to decompose Madagascar's rural-urban inequality into disparities in household and community attributes, circumstances, and assets; as well as differential returns to these assets. Using the same technique, it then decomposes changes in per capita consumption between 2005 and 2010 into explanatory factors by quintile of the consumption distribution. In "Flexible Poverty Profiling and Prediction of the Severity of Poverty in Madagascar" (McBride and Osborne 2016), the authors use "machine learning" algorithms to profile and predict levels of welfare in a manner that allows the data to iteratively shape the prediction model and isolate the most important predictive variables. The paper "Labor Demand Estimation in Rural Madagascar: Shadow Wages and Allocative Inefficiency" (Jodlowski 2016) uses multistage econometric estimation to analyze the determinants of labor demand by rural households, both in their agricultural and off-farm activities. Finally, in "Transactions Costs, Poverty, and Low Productivity Traps: Evidence from Madagascar's Informal Micro-Enterprise Sector," Bi and Osborne (2016) analyze the performance of urban-based, informal owner-operated microenterprises with respect to productivity and employment creation, using econometric methods that account for possible selection bias. The main insights of these papers and their policy implications are collected here.

Although conditions have been extremely unfavorable for poverty reduction—with real per capita GDP declining between 2001 and 2012—the poverty headcount rate has stabilized at approximately its 2001 level.[4] Households were buffeted by a variety of climatic and economic shocks, but the poor have adopted flexible strategies to return their living standards to previous levels.

Given the nature of macroeconomic and political events over the period, urban poverty rates fluctuated more widely than (and sometimes in opposite directions to) rural ones. Between 2001 and 2010, the national poverty rate moved in line with the urban headcount poverty rate. Between 2001 and 2005, both rose, but then they fell again in 2010, even as the rural poverty rate rose (Table I.1). This pattern no longer obtained between 2010 and 2012, however, when reductions in rural poverty offset increased poverty in urban areas to produce a slight decline in the national poverty rate to its 2001 level.

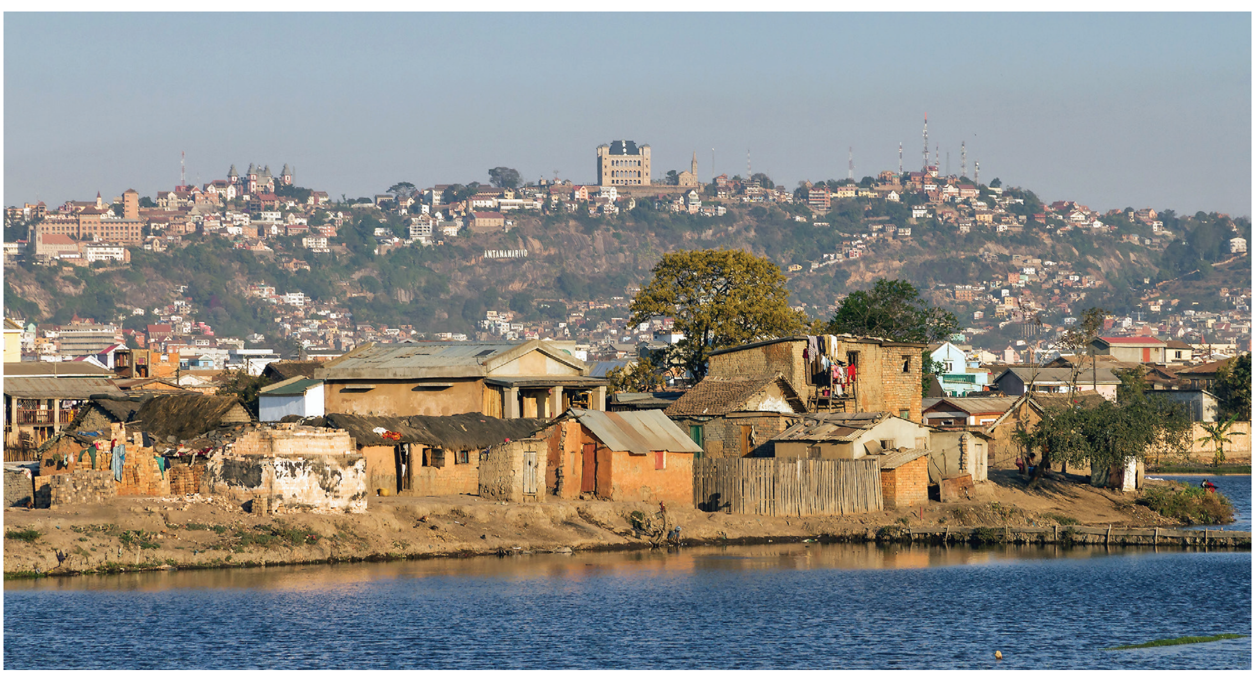
With approximately 78 to 80 percent of the rural population remaining poor throughout the period, perhaps a more meaningful indicator of rural poverty is the poverty gap index, which measures the severity of poverty. Over the years 2001 to 2012, this index moved in opposite directions in rural and urban areas. The national poverty gap index finished lower in 2012: on average, the poor lived on 32.2 percent less than the poverty line, relative to 35.9 percent less in 2001 (Belghith, Randriankolona, and Osborne 2016).

TABLE I.1: Trends in Poverty and Inequality (National Basic Needs Poverty Rates)

	2001	2005	2010	2012
Poverty gap index (mean percentage shortfall of poverty line)				
Urban	11.8	13.6	8.9	11.8
Rural	40.5	34.8	36.7	36.4
Total	35.9	31.3	32.0	32.2
Headcount poverty rate (percentage of the population)				
Urban	34.1	40.8	29.8	35.5
Rural	77.7	79.6	80.1	77.9
Total	70.8	73.2	71.7	70.7
Inequality indicators				
Gini Coefficient	46.9	38.9	42.7	41.0
P90/P10	8.13	4.96	6.01	6.32

Source: Belghith, Randriankolona, and Osborne 2016, using Enquête Périodique auprès des Ménages (EPM) and Enquête Nationale sur les Objectifs Millenaire du Développement (ENSOMD).

Over the period 2001–12, the population responded to fluctuations in returns to their income-generating activities by shifting effort and resources into and out of agriculture, services, and manufacturing sectors. As the returns in urban-based sectors fell in 2005, employment shifted into agriculture; when in 2010 returns in agriculture fell, labor shifted into nonfarm enterprises, and primary and secondary employment in services in particular rose (Belghith, Randriankolona, and Osborne 2016). The poor accumulated assets, including more education and transportation assets (Thiebaud, Osborne, and Belghith 2016). However, these strategies could not fully offset the weak demand for labor. In 2010, those

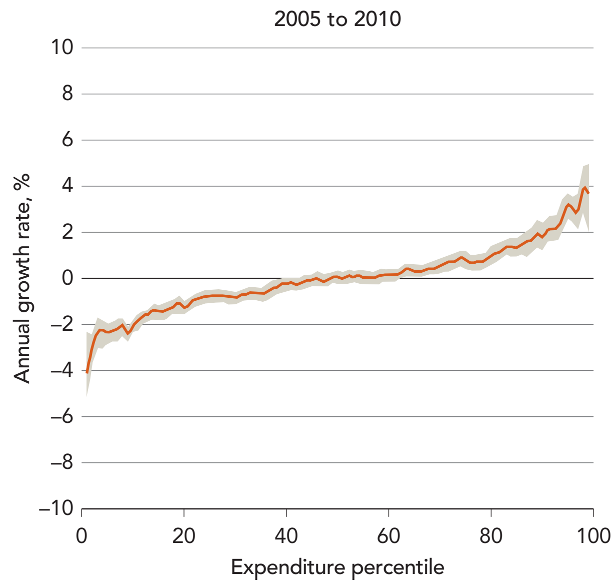


seeking but not finding work increased, even as secondary employment rose. Wages increased just slightly, and only for male workers in 2010, but then returned to their former (2005) levels (Belghith, Randriankolona, and Osborne 2016).

Inequality in Madagascar fluctuates significantly over time in response to climatic and price shocks. Particularly severe weather shocks in 2010 resulted in a decline in well-being for those at the bottom of the consumption distribution in that year vis-à-vis 2005. In combination with increasing returns to urban-based work relative to those in 2005, this led to a regressive growth pattern and increased inequality (Figure I.3). Although this implies significant consumption risks and welfare losses, one cannot assess persistent inequality (inequality in households' lifetime living standards) without tracking households over time (with panel data.) Madagascar's level of inequality as measured by the ratio of consumption for the top decile to that of the bottom decile (P90/P10) ranged from 5 to 8 over the period—a low level relative to the 13.4 average for low-income countries.

An important explanation for Madagascar's persistent poverty is its lack of progress in generating remunerative employment in the nonagricultural and urban sectors. As shown in Thiebaud, Osborne, and Belghith (2016), the returns to education and work are higher in urban areas. In addition, key attributes of urban communities—in particular greater access to markets, health centers, and other services—increase welfare (all else equal). In addition, rural households have been more adversely impacted by climatic risks. Between 2005 and 2010, the returns to economic activities in rural areas fell significantly for all quintiles of the consumption distribution. This, plus climatic shocks and to a lesser extent health shocks, explain the decline in welfare in the bottom two quintiles over these two years captured in figure I.4. Figure I.5 shows the main determinants of consumption changes by quintile and their direction of influence. As households responded by seeking greater off-farm employment, as shown, male-headed households were more successful than female-headed ones in offsetting these losses. This was because they were better able to secure employment in services sectors with apparently higher profitability, whereas females were more likely to find such employment in the primary and industrial (light manufacturing) sectors. Yet, over the period

FIGURE I.4: Incidence of Consumption Growth (Total)

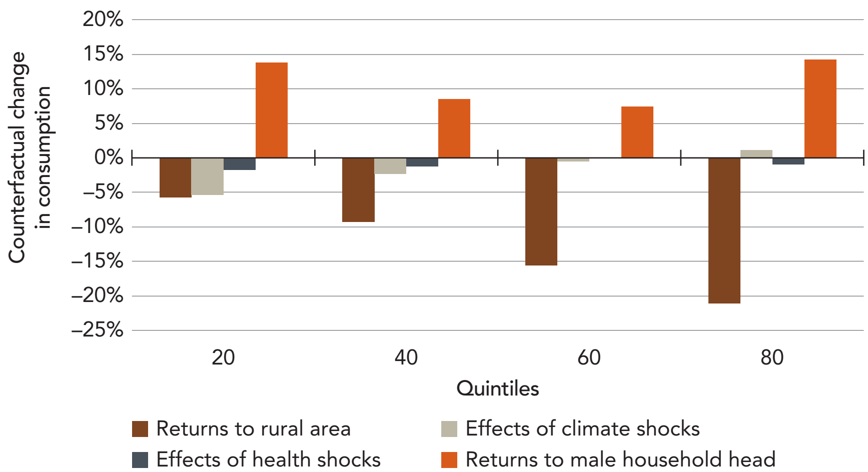


Source: Calculations using EPM 2005, 2010, reported in Belghith, Randriankolona, and Osborne 2016.

2005–10, despite having accumulated more assets, the rural population was unable to completely offset a decline in the returns to agriculture through entry into off-farm work. Returns to land fell by 6 percent—further for the poor—and health shocks compounded the toll (figure I.6).[5]

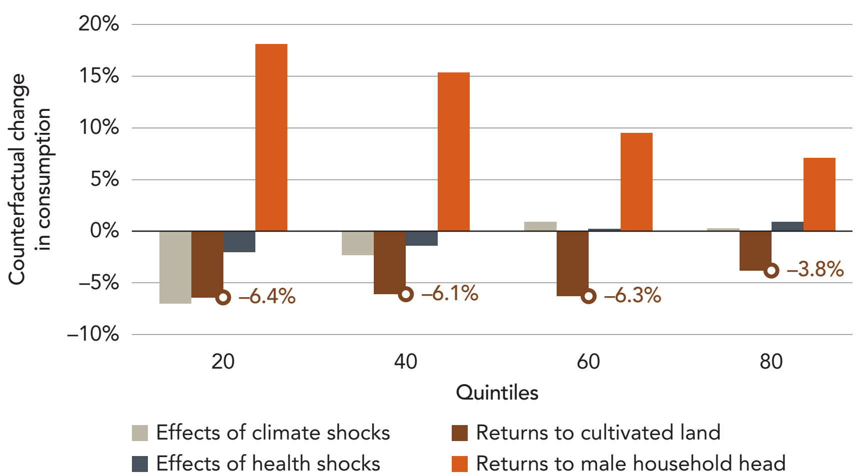
The key factors reducing agricultural incomes in 2010 were domestic rice policies and deteriorating transport conditions, which weakened internal market integration. In the face of rising world prices for rice, Madagascar's overwhelmingly dominant crop and staple food, in 2007 the government removed tariffs on rice imports and decreased *ad valorem* taxes, then in July of 2008 removed the value-added tax on rice imports completely. Anticipating drought and further increases in the world price, the government of Madagascar preordered rice imports (50,000 metric tons of Indian rice) and banned rice exports. These measures kept the price of rice relatively stable for consumers, yet producers were unable to benefit from rising world prices. In addition, rising transport costs reduced rural earnings. Between 2005 and 2010, the average real price to transport a 50 kilogram bag of rice rose 42 percent—from $US1.40 in 2005 to $US2.00 (2005 dollars), and to a higher level—$2.20 for the lowest consumption quintile of the population, using Enquête Périodique auprès des

FIGURE I.5: Main Determinants of Change in Consumption (2005–10)



Source: Thiebaud, Osborne, and Belghith 2016.
Note: Effects smaller than 2% or not significant for the bottom quintile are not pictured.

FIGURE I.6: Main Determinants of Change in Consumption (Rural Households, 2005–10)



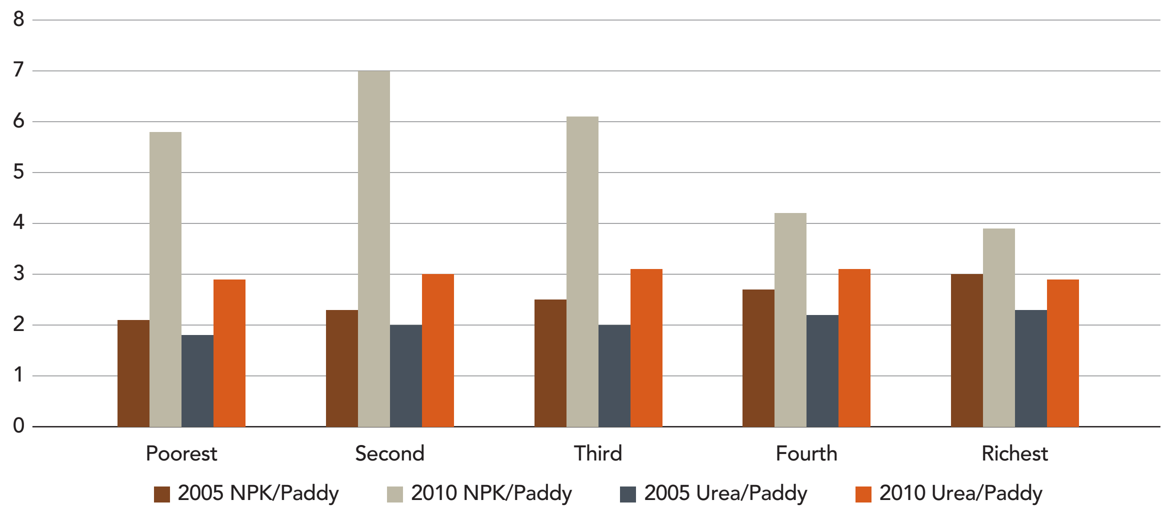
Source: Thiebaud, Osborne, and Belghith 2016.

Ménages (EPM) 2005, 2010. As a result, between 2005 and 2010, the ratio of paddy prices to fertilizer dropped precipitously and was more closely correlated with consumption than before (figure I.7).

These findings are supported further by a flexible profile of the severity of poverty (McBride and Osborne 2016). Of the many available household, regional, and community-level variables which one might expect to be correlated with welfare, those that are most predictive of more severe poverty are the following, in order of importance:[6]

1. Living in a community with levels of electrification less than 27 percent of households
2. Having a non-university-educated household head (Having a university education makes it very likely to have higher incomes in urban areas, as would be expected.)

FIGURE I.7: Average Nominal Price Received for Rice Paddy (by Consumption Quintile)



Sources: EPM 2005 and EPM 2010.

3. Having an illiterate head of household (Other distinctions in educational attainment have little predictive power.)
4. Living in greater remoteness from the nearest major urban center (This variable predicts welfare better than other measures of access to services.)
5. Receiving lower prices for paddy rice
6. Having lower livestock holdings

For agricultural households analyzed separately, the key predictive variables in order of importance are the following:

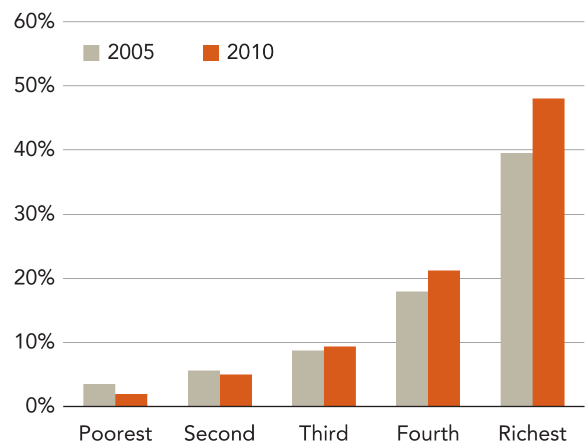
1. Lower cultivated land
2. Remoteness from the nearest major urban center
3. Living in a community with lower levels of electrification
4. Receiving a higher percentage of revenues from agriculture
5. Receiving a lower price for paddy rice

The combined results of Thiebaud, Osborne, and Belghith (2016) and McBride and Osborne (2016) provide evidence that, while the effects of rice prices are always somewhat heterogeneous within a population, low rice prices in Madagascar increase poverty in rural areas. An inter-temporal decomposition of changes in consumption (Thiebaud, Osborne, and Belghith 2016) shows that declines in the returns to land are strongly associated with more severe poverty, and in 2010 the households facing lower rice prices have lower consumption. Thus, the benefits to poor net consumers are more than offset by the decline in the incomes of poor rice producers. Both papers show that this is the case even at the bottom of the distribution.

Our findings on the role of electricity (McBride and Osborne 2016; Thiebaud, Osborne, and Belghith 2016) provide indicative but not conclusive evidence on its importance for alleviating poverty. Electrification may simply proxy for community-level wealth or economic activity, the effects of which are difficult to disentangle from those of electricity provision. Nonetheless, the combined results are suggestive of a positive causal effect on incomes. First, we consider the possibility that electricity is merely a proxy for other urban attributes. We examine the correlation between the level of a community's electrification and an indicator for urban area, regional indicators, and all other indicators of remoteness which would be expected to be correlated with urban agglomeration. We find that even controlling for these variables, the degree of electrification varies substantially, within and across rural and urban areas of varying degrees of remoteness and access to services (McBride and Osborne 2016).[7] Electrification was a more powerful predictor of welfare than any other of the available indicators of spatial advantage or

economic density. Moreover, Thiebaud, Osborne, and Belghith (2016) show that between 2005 and 2010, increased access to electricity was associated with a small but positive and statistically significant change in consumption for parts of the consumption distribution. Over the period the percentage of households with electricity increased only slightly, from 15 percent to 17 percent, but the correlation between electrification and consumption increased. Electrification reached 0.5 percent more households in the third quintile of consumption in 2010 consumption than it did in 2005, 3.2 percent more of the fourth quintile, and 8.5 percent of the top quintile, and less of the bottom two (figure I.8). Since the ability to shift into off-farm work was a key strategy for coping with poor returns to agriculture in 2010, communities with better access to electricity were likely better able to support more productive nonfarm enterprises (NFEs). Jodlowski's (2016) finding that greater electrification was significantly correlated with NFE revenues in 2010 (but not in other years) supports this hypothesis. Electricity can raise incomes where there is potential for NFEs and a certain level of demand, but there is no evidence available that it can do so in the remotest and poorest areas.[8]

FIGURE I.8: Percent of Households with Electricity in Community (by Consumption Quintiles, 2005 and 2010)



Note: Data are not representative at the community level.
Sources: EPM 2005, 2010.

In addition, our findings underscore the importance of reducing transport costs for poverty reduction. First, in 2001 and 2005, higher transport costs were associated with lower levels of rural NFE revenue (Jodlowski 2016). Although similar effects are no longer evidenced in 2010, there is a close association between higher transport costs, worsening terms of trade in agriculture, and declining consumption (Belghith, Randriankolona, and Osborne 2016; Thiebaud, Osborne, and Belghith 2016). Moreover, the time to reach urban centers and health centers is not correlated only with poverty but is highly predictive of severe poverty (McBride and Osborne 2016). These results are unlikely to simply reflect a migration of poorer people to more remote areas. Rather, the political crisis of 2009 reduced the

availability of funding for road maintenance at a time when oil prices were high relative to 2005.[9]

More productive and remunerative off-farm employment is the primary route out of poverty, in Madagascar as well as in other poor agricultural economies, but it requires that constraints to larger, more efficient enterprises be alleviated. An examination of rural labor markets (Jodlowski 2016) as well as a study of urban-based informal enterprises (Bi and Osborne 2016) suggest that the informal microenterprise structures prevailing in Madagascar result in a major misallocation of resources in the economy. While such enterprises provide a means to generate a livelihood, the productivity and income losses resulting from this structure are substantial. As they do in many other poor countries where informal microenterprises have been studied, they do not tend to scale up and employ more workers over time. As Jodlowski (2016) and Bi and Osborne (2016) show, they underemploy workers. Therefore, as the primary source of off-farm employment, this configuration of production reduces employment and wages, making the rural-urban transition much more difficult.

Despite the flexible coping strategies Madagascar's population exhibited over the years covered by this report (2001–12), as currently configured, the potential for Madagascar's rural labor markets to generate more productive employment remains low. According to estimates presented in Jodlowski (2016), rural labor markets, where households are the primary employers, are subject to considerable frictions. This results in low demand and a low willingness to pay a (shadow) wage, whether to employ labor on the farm or in NFEs. Although it is not possible with the available data to identify the precise source of the frictions, they likely represent some combination of the risks and (nonfinancial) costs of identifying, training, supervising, and releasing workers (Jodlowski 2016). Thus, the source and magnitude of these costs may differ by community and by household attributes. They may also vary between farm and off-farm enterprises for the same household. On average, Jodlowski finds the estimated size of the friction to be greater in the on-farm sector than in NFEs.[10] At the same time, the demand for on-farm labor is relatively responsive to the shadow wage, whereas NFE labor is unresponsive. This suggests that in agriculture labor input is easier to adjust as needed, as profitability conditions change. Agricultural laborers may also be more easily substituted for each other, as the tasks performed are less complex or specialized. For NFEs, however, there appear to be greater rigidities involved in adjusting labor inputs, which deters these enterprises from hiring more workers. To the extent that labor input is adjusted, the main flexibility is in the hours of existing workers rather than the number of different employees hired. NFE labor may require more effort to find, train, and supervise. More skilled or specialized workers may also expect a more regular contract, even if this contract is informal, and this implies greater risk to the employer. Thus, rural nonfarm entrepreneurs prefer to accept lower expected profits than to incur these costs. The primary vehicle of employment by household-owned and -operated NFEs, therefore, is through self-employment, with minimal potential for generating employment for other workers.

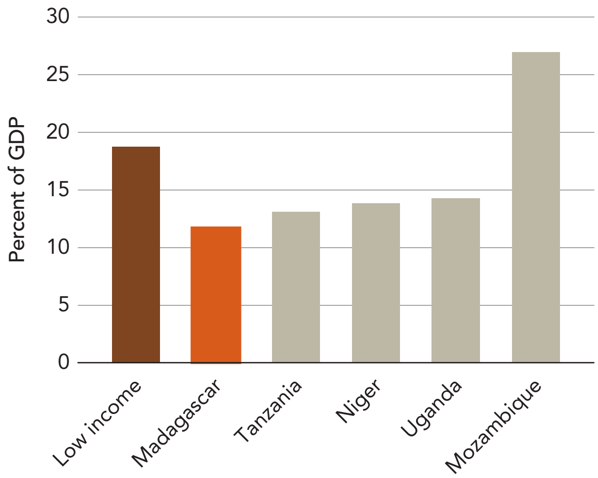
The potential for remedying labor market frictions is unclear. The primary source of these frictions appears to be market failures related to risks and the challenge of incentivizing workers to be productive, honest employees ("agents") acting largely in accordance with the objectives as the owner of the enterprise (the "principle"). These principle-agent problems increase the transactions costs of employing workers and cause market failures that household enterprises overcome by employing family members, first and foremost, and, in the vast majority of cases, by staying small. Because the benefits of incurring these transactions costs must exceed their costs in order for enterprises to hire more workers, where the marginal returns to labor are low, transaction costs may loom too large as a percentage of these gains. Thus, in principle, actions that would improve the profitability of enterprises, such as investments in public infrastructure, would translate into jobs. At present, however, there is no compelling evidence (from Jodlowski) that improving the profitability of rural NFEs in this manner would result in a significant increase in either wages or employment by NFEs.

Similarly, the potential for urban-based, microenterprises to generate greater employment and economic gains is limited. Using detailed 2012 data on informal, owner-operated microenterprises (OOMEs) in Madagascar's cities and towns, Bi and Osborne (2016) assess the potential of these enterprises to achieve higher incomes for their owners and offer remunerative employment to workers. For both single-worker OOMEs—those which employ only their owner—and multiworker OOMEs, which employ family, other unpaid, as well as paid labor, the activities pursued range from logging and mining to household services, transport services, and light manufacturing. Seventy percent of OOMEs are the single-worker variety, and this variety has significantly lower returns to capital and to the owner's labor due to unexploited "profit" economies of scale.[11] The wage penalty of owning and operating a single-worker OOME rather than working for others (controlling for worker ability, characteristics, sector, and location of employment) is approximately 60 percent of the mean wage. Owners of multiworker OOMEs, however, earn a "wage" premium of approximately 68 percent of the mean wage, controlling for individual characteristics (excluding returns to capital).

The persistent prevalence of microenterprises that are too small to be productive can be explained by market failures—high transactions costs and risks—that are particularly difficult to overcome in a poor economy. OOMEs (correctly) perceive a lack of demand for their products to be their most immediate constraint, but it would be more efficient for there to be fewer, larger enterprises serving the same level of demand than numerous OOMEs. Given increasing returns to capital, OOMEs could in principle increase their profitability by expanding a little at a time, reinvesting growing profits and growing to a more efficient and profitable scale. To explain their lack of growth, therefore, requires a combination of conditions.

One issue is the lack of entry by larger, more efficient, typically formal firms, which would provoke a restructuring of the market and draw workers into more remunerative work. In addition, an economy characterized by poor entrepreneurs (the OOMEs) inhibits their growth. The marginal utility of consumption for entrepreneurs' poor families is high, while returns to investing small incremental amounts in tiny enterprises are low. In the presence of increasing returns, firms must be created larger or grow rapidly to enjoy those returns. Thus, constrained to consume all of their income, entrepreneurs lacking sufficient external financing cannot grow their businesses. Breaking out of this low-productivity equilibrium (or "trap") would require a more substantial increase in scale than poor households can afford. At the same time, transactions costs associated with external financing are high. Due to the difficulties associated with monitoring the use of firm resources, whether by creditors or potential partners (another principle-agent problem), the transactions costs of credit, partnership arrangements, and share capital are high. Microcredit, if available, must come with interest rates adequate to cover the costs of screening and enforcement of repayment, and these costs are higher on a per-dollar (or ariary) amount for small loans. Similarly, partnerships do not form for the purposes of expansion precisely because entrepreneurs' level of investible capital is low relative to these transactions costs. Finally, the frictions associated with employing and incentivizing workers further hinder firms' profitability and growth.

FIGURE I.9: Credit to the Private Sector as a Percent of GDP, Comparison Countries and Low Income Average (Average 2011–14)



Source: WDI.

The tentative policy implications of these findings are as follows: First, it would likely have little effect to encourage informal microenterprises to simply register without taking additional steps to improve the credibility of their financial statements and integrity of firm resource uses. Rather, levers to strengthen the information environment (through adoption and verification of accounting practices, credit reporting, and other means) would be needed to reduce the transactions costs for potential creditors and partners. In addition, while microloans to the tiniest, single-worker OOMEs may help to provide employment for their poor owners, they would have little impact on overall productivity, employment, and wage growth.

Ultimately, significantly reducing the misallocation of capital and labor in Madagascar's economy would require a steadily growing presence of larger, more formal firms that compete for markets (Aghion, Akcigit, and Howitt 2013). At the same time, a significant improvement in productivity could be attained through the alleviation of constraints facing OOMEs that have already achieved a certain scale, that employ workers, and that demonstrate basic entrepreneurial skills. Given the presence of increasing returns, such firms could invest more and hire more workers if they had access to external financing. A full 92.5 percent of OOMEs received their assets via a gift or inheritance or created them with their own savings, and only 1.2 percent of them created them through some type of loan, most of which were informal. Thus, efforts to speed the development of a financial sector capable of allocating financial savings to the most promising investments could be effective in stimulating more dynamic change. Madagascar's financial markets are underdeveloped relative to other low-income and SSA countries, with a low level of credit to the private sector as a percentage of GDP (figure I.9). One key step would be to improve the monitoring and enforcement environment for credit, partnerships, associations, and corporate investors. In 2012, for example, credit registries and bureaus were almost nonexistent in the country, and the strength of legal rights to enforce repayment rated only 2 out of a 10-point scale in Doing Business's *Getting Credit*.[12] Finally, depending on the status of existing financial institutions, policy makers could consider enhanced liquidity or risk mitigation measures to expand the capacity to serve larger "micro" and small and medium-sized enterprises.[13]

This series of papers also provides insights on gender-related disparities in opportunities in Madagascar. Although female-headed households are not consistently poorer than male-headed ones (McBride and Osborne 2016), men earn significantly higher wages than women (Belghith, Randriankolona, and Osborne 2016). When educational attainment, region, and urban milieu are considered, men earned 37 percent more than women in the labor market in 2012 (Bi and Osborne 2016). Female entrepreneurs are less likely to own and operate a multiworker microenterprise and more likely than men to own and operate their less profitable single-worker versions. Among single-worker firms, men earn higher profits, all else equal (Bi and Osborne 2016), and appear to face fewer obstacles in undertaking certain economic activities than women, a disparity in access to opportunities that widened substantially in 2010, as shown in Thiebaud, Osborne, and Belghith (2016).

These findings raise additional issues for further investigation: In particular, further investigation into the sources of labor market frictions, possible mitigating factors, and policy levers would be beneficial, both for formal and informal job creation. It would be worthwhile exploring further the constraints faced by female entrepreneurs, as well as collecting higher quality rural farm and NFE data, perhaps detached from the EPM. Finally, as infrastructure improvements are made, it would be extremely valuable to evaluate the impacts of them on incomes and well-being in a rigorous manner.

NOTES

1. Madagascar ranks fifth in Sub-Saharan Africa in renewable water resources per capita, according to WDI data.
2. Based only on countries with data.
3. Madagascar's AGOA preferences were reinstated in 2014.
4. Because poverty headcount rates moved in opposite directions to growth in two of the three subperiods (in particular, 2005–10 and 2010–12), the elasticity of poverty with respect to growth would have the wrong sign for these and the full period, and it is not considered an informative measure of the upside potential of positive growth to reduce poverty in the country.
5. It is not clear to what extent adverse health shocks caused greater poverty versus the decline in incomes causing adverse health.
6. This analysis excludes household size and composition variables, which overstate the adverse welfare effect of large households with more children using the per capita income welfare indicator.
7. A multivariate regression of electrification on the full set of available geographic and remoteness variables leaves 42 percent of this variation unexplained.
8. The "returns" to electricity were estimated to have fallen for the bottom quintile between 2005 and 2010 (Thiebaud, Osborne, and Belghith 2016).
9. Average world oil prices were 48 percent higher in 2010 than 2005. See World Bank, "Commodity Markets," database, http://www.worldbank.org/en/research/commodity-markets.
10. These differences may not be statistically significant or stable over time, as the farm-based estimates are for one year only.
11. Since returns are measured in terms of profits rather than decreasing costs, they represent the combined effects of declining average costs and increased market power.
12. Although these indicators had improved by 2016, they still show weak performance.
13. Any interventions would ideally be designed to ensure that (i) the information and monitoring environments were also improving, (ii) access to assistance was competitive and fairly distributed among individual and group enterprises, and (iii) subsidies did not undercut other developments in credit markets.

REFERENCES

Aghion, P., U. Akcigit, and P. Howitt. 2013. "What Do We Learn from Schumpeterian Growth Theory?" Schumpter Lecture presented at the Swedish Entrepreneurship Forum, Stockhim, January 2013. http://scholar.harvard.edu/files/aghion/files/what_do_we_learn_0.pdf?m=1361377935.

Belghith, N., P. Randriankolona, and T. Osborne. 2016. "Madagascar Poverty and Inequality Update: Recent Trends in Welfare, Employment, and Vulnerability." Washington, DC: World Bank.

Bi, C., and T. Osborne. 2016. "Transactions Costs, Poverty, and Low Productivity Traps: Evidence from Madagascar's Informal Micro-Enterprise Sector." Washington, DC: World Bank.

Jodlowski, M. 2016. "Labor Demand Estimation in Rural Madagascar: Shadow Wages and Allocative Inefficiency." Washington, DC: Cornell University and the World Bank.

McBride, L., and T. Osborne. 2016. "Flexible Poverty Profiling and Prediction of the Severity of Poverty in Madagascar." Washington, DC: Cornell University and the World Bank.

Thiebaud, A., T. Osborne, and N. Belghith. 2016. "Isolation, Crisis, and Vulnerability: A Decomposition Analysis of Inequality and Deepening Poverty in Madagascar (2005–2010). Washington, DC: World Bank.

World Bank. 2014. *Face of Poverty in Madagascar: Poverty, Gender, and Inequality Assessment*. Washington, DC: World Bank.

CHAPTER

1

Madagascar Poverty and Inequality Update: Recent Trends in Welfare, Employment, and Vulnerability

Nadia Belhaj Hassine Belghith
Patrick Leon Randriankolona
Theresa Osborne
June 2016
The authors thank Rachel Wang and Alessia Thiebaud for capable research assistance and Kristen Himelein for helpful advice.

Summary

For Madagascar, the years 2001–12 lacked both economic growth and progress in alleviating poverty. Political, economic, and climatic shocks caused fluctuations in poverty, producing an increase in headcount poverty from 2001 to 2005, followed by a modest decline for several years, and a rise in 2012 when the headcount poverty rate stood at 70.7 percent, using the national poverty line, and essentially the same rate as in 2001 (70.8). From 2001 to 2013, perhaps somewhat surprisingly, increases in the headcount poverty rate were accompanied by decreases in the severity of poverty and inequality and vice versa. Households living near the poverty line—that is, those which are less poor than 70 percent of the population—are buffeted by conditions in urban settings, whereas the depth of poverty for the poorest of the poor is affected to a greater extent by conditions in rural areas.

Updated Poverty Statistics

Despite adverse conditions over the past decade, trends in poverty headcount rates in Madagascar have stabilized, albeit at a high rate, approximately 71 percent.[1] Table 1.1 shows updated estimates of the national poverty rate in years with consumption data available: 2001, 2005, 2010, and 2012. As real per capita gross domestic product (GDP) declined from US$294 to US$267 (in

TABLE 1.1: World Bank Revised Headcount Poverty Estimates (National Basic Needs Poverty Line), Earlier Estimates, and Real GDP per Capita in Years with Consumption Data

Year	2001	2005	2010	2012
Official poverty estimates	69.7%	68.7%	76.5%	71.2%
Percent of population in poverty, earlier estimates (World Bank 2014)	70.8%	75.0%	75.3%	n.a.
Total (percent of population) in absolute poverty, final revised	**70.8%**	**73.2%**	**71.7%**	**70.7%**
GDP per capita in 2005 U.S. dollars	$294.0	$275.5	$273.2	$267.2

Sources: Bank staff using Enquête Périodique auprès les Ménages (EPM), Enquête Nationale sur les Objectifs Millenaire du Développement (ENSOMD), and World Development Indicators (WDI).
Note: Poverty line is estimated using 2010 EPM survey and adjusted for inflation in each year.

2005 U.S. dollar purchasing power parity, PPP), poverty headcount rates based on the basic needs approach rose from 70.8 percent in 2001 to 73.2 percent in 2005, then fell slightly to 71.7 percent in 2010 and to 70.7 percent in 2012—a return to their 2001 level.

These estimates vary from those published earlier and indicate that the declining trend observed in the headcount poverty rate as of 2012 (INSTAT 2014) began earlier than previously thought. As with any poverty measurement, the precise methods used to estimate the welfare aggregate (typically consumption) and the poverty line can have an impact on levels and, in some cases, trends. Headcount poverty estimation is sensitive to small changes in the welfare indicator or the poverty line, and the adjusted figures reported here are within the confidence interval for earlier reported estimates.[2]

Nonetheless, the revised estimates are likely to be more exact as they more accurately reflect the best available information on the geographic structure of the Malagasy population. In particular, the weights used by Madagascar's National Institute of Statistics (INSTAT) to compute population-level statistics implied a spatial partition of the population, which deviated from the best estimates of its actual partition.[3] In effect, too low a weight had been assigned to urban households and too high a weight to rural households; since urban households exhibit lower rates of poverty, this tended to overstate national poverty rates and underestimate consumption growth between 2005 and 2010.[4] (See annex 1A for more details on this and other methodological issues.)

Regardless of these adjustments, Madagascar's poverty rates are exceedingly high, and according to internationally comparable estimates are the highest in the world.[5] Using the World Bank's international poverty lines of US$1.90 per capita per day (in 2011 PPP), poverty in Madagascar is 77.8 percent (table 1.2).[6]

Close to 80 percent of Madagascar's population lives in rural areas, and rural poverty rates are more than twice as high as urban rates. As shown in table 1.3, although rural poverty rates have stayed fairly flat— having risen slightly after 2001, then fallen back to 2001 levels in 2012—urban poverty rates have fluctuated much more, from 34 percent in 2001 to over 40 percent in 2005, 29.8 percent in 2010, and once again close to the 2001 level in 2012.[7]

TABLE 1.2: Poverty Rates Using International Poverty Lines, Povcalnet Method (Corrected Sampling Weights, no Application of Regional Deflation)

Year	2001	2005	2010	2012
US$1.90 2011 PPP	**68.7**	**74.1**	**81.8**	**77.8**
US$3.10 2011 PPP	**84.1**	**89.9**	**92.9**	**90.5**
US$1.25 2005 PPP	76.7	80.7	84.8	83.9
US$2.00 2005 PPP	88.2	92.1	93.6	93.3

Sources: EPM and ENSOMD.
Note: PPP = purchasing power parity.

TABLE 1.3: Trends in the Poverty Gap (Mean Percentage Shortfall of Consumption Relative to Poverty Line)

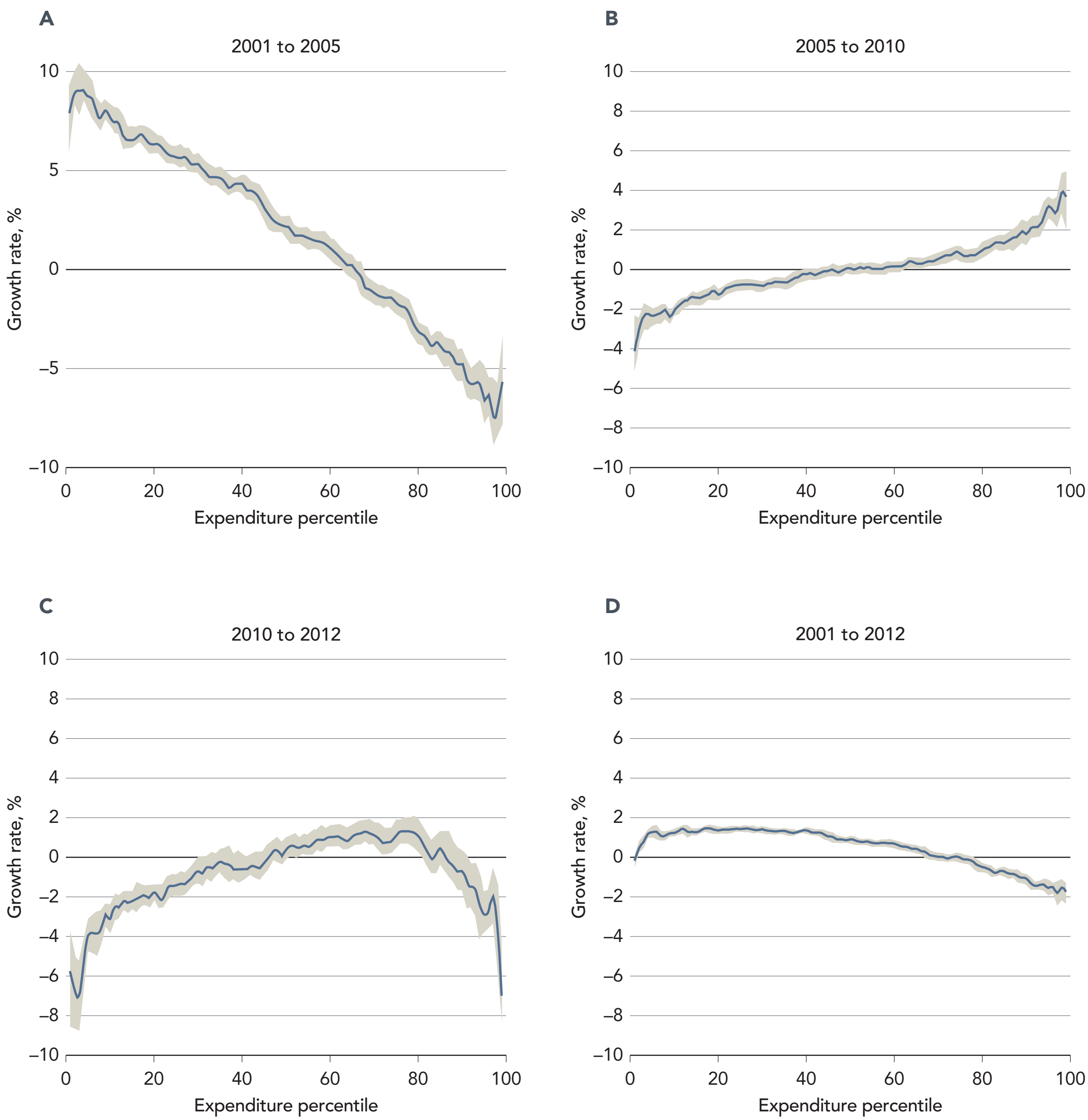
	2001	2005	2010	2012	Change 2010–2012
Urban	11.8	13.6	8.9	11.8	2.9
Rural	40.5	34.8	36.7	36.4	–0.3
Total	35.9	31.3	32.0	32.2	0.2

Sources: EPM and ENSOMD.
Note: Poverty line = Basic needs, World Bank.

The Distribution of Growth and Changes in Inequality

Fluctuations in the headcount poverty rate from 2001 to 2012 mask significant changes in the distribution of consumption growth. In fact, over this period, poverty headcount rates have risen in years when growth has been more progressive, and vice versa: poverty headcount rates have fluctuated with the fortunes of households living close to the poverty line, whereas those falling lower (and higher) in the distribution have been impacted quite differently. Between 2001 and 2005, when a significant percentage of the nonpoor population fell into poverty, mean consumption levels of the poor nonetheless largely improved. This is shown in the growth incidence curves (figure 1.1). However, as shown this pattern reversed after 2005, with declines in real consumption below the 40th percentile and gains at the top. Between 2010 and 2012, the pattern is mixed, with declines at the bottom and top and improvements in the middle of the distribution—where there are

FIGURE 1.1: Growth Incidence Curves, 2001–12 and Subperiods (Total Percentage Changes)



still many poor people. Overall, when comparing 2012 to 2001, the bottom range within the poor population showed net gains in consumption: The regressive pattern of consumption growth after 2005 did not completely offset the gains made at the bottom of the distribution from 2001 to 2005.[8] The poverty gap, a measure of the severity of poverty, correspondingly fell, from 35.9 in 2001 to 31.3 in 2005, and inched up only slightly in 2010 to 32.0 and in 2012 to 32.2, still lower than in 2001, despite the decrease in real per capita GDP over the period (table 1.3).[9]

Despite negative (real) per capita GDP growth over the period 2010–12, the headcount poverty rate dropped by 1 percentage point due to a favorable distribution of

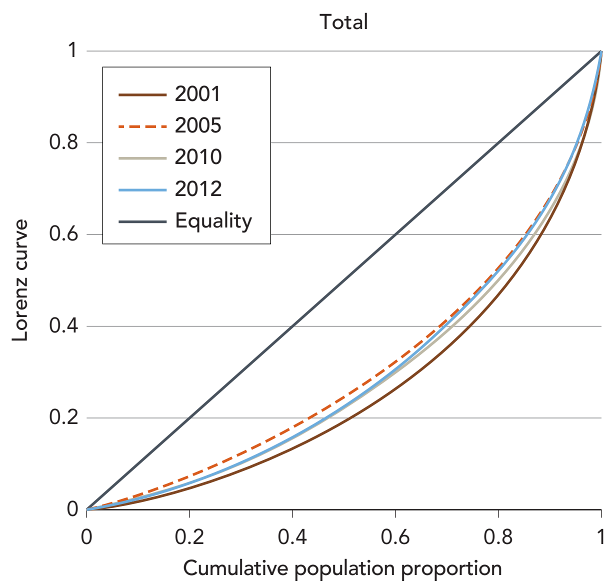
TABLE 1.4: Decomposition of Growth versus Inequality Contributions to Changes in Poverty Rates (2010–12)

	2010	2012	Actual change	Growth	Redistribution	Interaction
Poverty line 5 Basic needs (World Bank)						
Total	71.65	70.74	–0.91	1.58	–2.77	0.28
Urban	29.82	35.52	5.71	6.83	0.38	–1.51
Rural	80.12	77.93	–2.19	0.46	–2.61	–0.04
Poverty line 5 Food poverty line (World Bank)						
Total	58.28	57.43	–0.85	2.33	–3.05	–0.13
Urban	18.35	22.67	4.32	5.09	–0.41	–0.35
Rural	66.36	64.52	–1.84	0.63	–2.50	0.03

growth near the poverty line. Negative growth would have added 1.58 percentage points to the poverty rate (table 1.4) and an astonishing 6.83 percentage points to the urban poverty were it not for the distributional effects. In rural areas, the distribution of growth was pro-poor overall between the two years, with the redistribution effect accounting for a reduction in the headcount poverty rate by 2.61 percentage points.

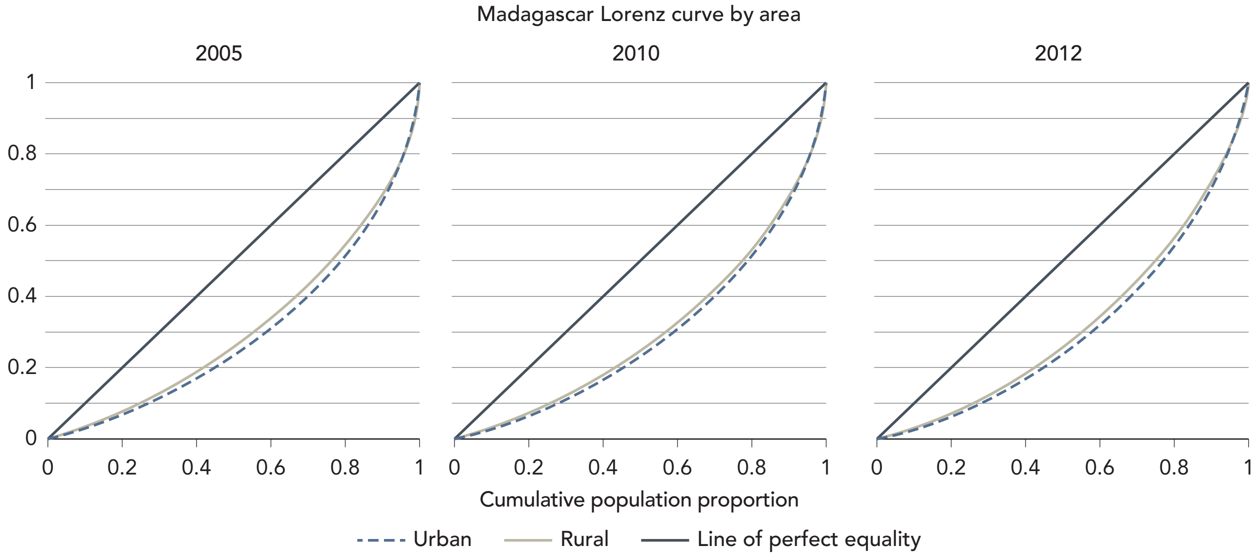
Whereas Madagascar's poverty rates are exceedingly high, inequality is in line with that of other low-income countries and has fallen over the period. Urban areas display more unequal distributions of real per capita consumption than the rural zones; however, inequality in rural areas increased over the period of analysis. Figure 1.2 shows the national Lorenz curves for all four survey years, again reflecting the generally equalizing trends between 2001 and 2005, which then partially reverse thereafter. Lorenz curves that depict the divergence from perfect equality at different parts of the distribution are shown in figure 1. 3.

FIGURE 1.2: Inequality (Lorenz Curves) for 2001, 2005, 2010, and 2012 (Cumulative Share of Welfare Accruing to x-axis Proportion of the Population)



The Gini coefficient—a summary, internationally comparable measure of inequality related to the Lorenz curve—also shows fluctuations, but finished the period lower.[10] Starting at a high of 46.7 in 2001, it was 41.0 in 2012, relative to a low income average of 40 for countries with available data over period 2007–11, according to the World Development Indicators (WDI). However, the Gini coefficient does not capture distributional changes that may occur in different parts of the welfare distribution. It would reflect a redistribution from the middle of the distribution toward the bottom in the same manner as a redistribution from the top to the middle, for example. Information on consumption shares by population quintiles—in particular the ratio of average consumption of the top decile of households to that of the bottom decile (P90/P10)—helps to overcome this shortcoming. These measures show that much of the increase in inequality after 2005 is driven by a decline in the welfare share accruing to the poorest segment of the population, which dropped from 6.97 to 5.89 percent (15 percent) between 2005 and 2012, except in the urban sectors where it declined by only 4 percent. These measures also reveal that in the case of Madagascar inequality does not reflect the presence of high wealth at the top of the distribution: As shown in table 1.5, the average consumption ratio for the top decile to the bottom decile has been between 5 and 8 over the period under study and thus is consistently lower than the low-income average over

FIGURE 1.3: Lorenz Curves and Inequality Coefficients



	2005				2010				2012			
		Income shares				Income shares				Income shares		
	Gini	p90/p10	Low. quintile	Top quintile	Gini	p90/p10	Low. quintile	Top quintile	Gini	p90/p10	Low. quintile	Top quintile
National	38.93	4.96	6.97	46.63	42.66	6.01	6.03	49.69	41.03	6.32	5.89	47.63
Rural	35.35	4.24	7.63	43.48	37.96	4.73	6.91	45.38	37.3	5.32	6.54	44.33
Urban	39.22	5.51	6.50	45.91	38.60	5.16	6.56	45.38	38.44	5.92	6.27	45.12

Sources: ENSMOD 2005, 2010, and 2012.

TABLE 1.5: Consumption (per Capita) Inequality Measures

Bottom half distribution		Upper half distribution		Interquartile range		Tails	
	p25/p10	p50/p25	p75/p50	p90/p75	p75/p25	p90/p10	Gini
2001	1.50	1.66	1.79	1.83	2.96	8.13	46.9
2005	1.39	1.44	1.53	1.61	2.22	4.96	38.9
2010	1.48	1.51	1.59	1.69	2.40	6.01	42.7
2012	1.53	1.56	1.62	1.63	2.52	6.32	41.0

Sources: EPM 2001–2010 and ENSOMD 2012.

2007–11 of 13.4 (per WDI). Finally, a single year's snapshot of the consumption distribution, particularly in a country such as Madagascar, which faces a high level of weather-related and other risk, can overstate inequality in households' lifetime welfare. A high level of interannual variation means that households are moving up and down in the distribution from year to year.

Trends in Agriculture and Employment

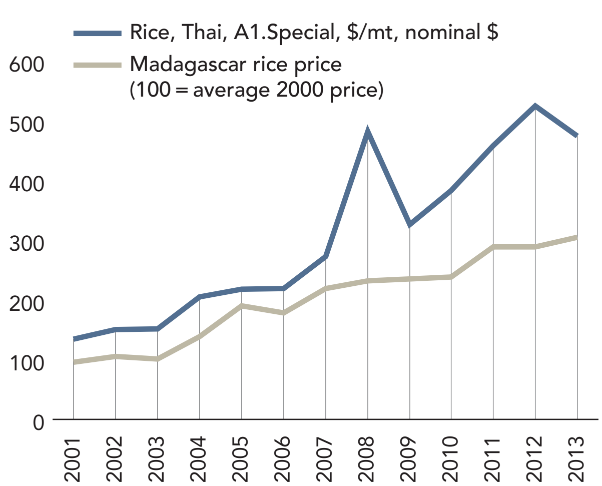
A combination of external and domestic shocks and policy responses buffeted the poor over the period 2001–12. Changes in the terms of trade in agriculture, consumer price inflation, weather shocks, and changing off-farm labor market conditions affected households throughout the consumption distribution. Households responded by adjusting their levels of secondary employment and self-employment, as well as their allocation of work across sectors as returns in various sectors shifted. Overall, workers reporting a wage saw little to no wage growth.

First, in 2001–2005, the political crisis of 2002 combined with cessation of most-favored-nation preferences to produce an adverse effect on labor markets. Urban employment in particular declined, and many people dependent for their incomes on the textiles sector fell into poverty. In addition, consumer price inflation, driven largely by local conditions, increased the cost of living. The consumer price index increased by over 18 percent in 2005 alone, with a cumulative 54 percent increase over the four years between 2001 and 2005. Although high inflation is not unusual for the country—annualized inflation has averaged more than 11 percent since 1965—over 77 percent of households reported an adverse effect of general price inflation in 2005, whereas only 2.9 percent did so in 2010 (and 0.8 percent in 2012). Because of these and other factors, the headcount poverty rate in urban areas increased dramatically, from 34.1 to 40.8 percent. At the same time, as figure 1.1 showed, consumption at the bottom of the distribution rose. Although it is unclear which conditions contributed most to this improvement, rising rice prices after 2005 may have improved the net incomes of rice producers (see figure 1.4).

Malagasy households responded to these shifts in circumstances by adjusting their labor supply and sectors of employment. The adverse labor market shocks in 2005 combined with reasonably favorable terms of trade in agriculture to induce a substantial movement of labor into agriculture. As shown in figure 1.5, one can discern a clear shift between 2001 and 2005 into agriculture for the top 40 percent of the distribution, with the remaining 60 percent maintaining their extremely high (over 80 percent) rates of primary employment in the sector. One also observes a decline in the percentage of household heads employed in manufacturing from 2001 to 2005.

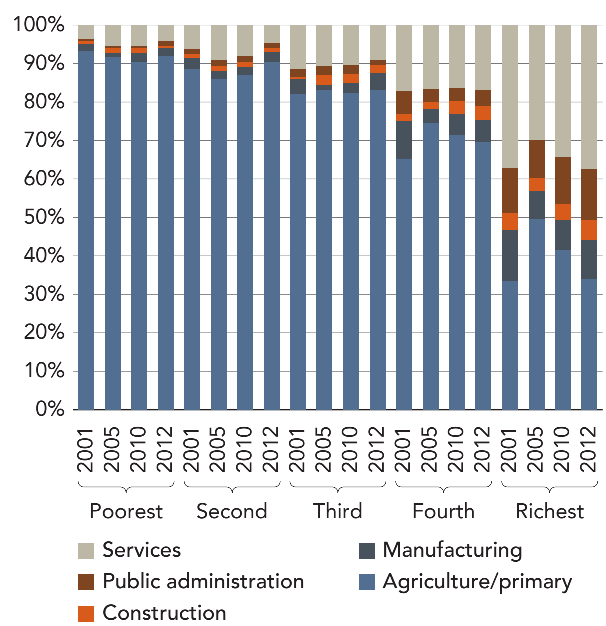
Between 2005 and 2010, the declining profitability of agriculture—in particular for rice cultivation—contributed to households' seeking employment outside of the sector. By 2010, although the world price of rice had continued its increasing trend, the terms of trade in agriculture shifted against producers, due in part to declining transport conditions and policies designed to maintain lower rice prices (see Thiebaud, Osborne and Belghith 2016). Between 2005 and 2010, the producer price of paddy rice fell (figure 1.6), despite generally rising world food prices. Moreover, in contrast to 2005, in 2010, households' consumption levels were positively correlated with paddy prices across quintiles. Moreover, as transport conditions deteriorated, for each percentage point increase in the time to reach input markets, the relative price of rice relative to

FIGURE 1.4: Rice Price Indices (2001 = 100)



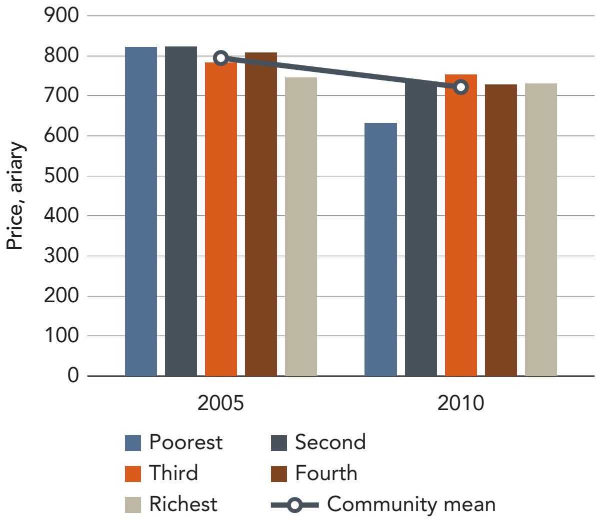
Source: FAO

FIGURE 1.5: Sector of Main Employment of Household Head (%) by Quintile and Year



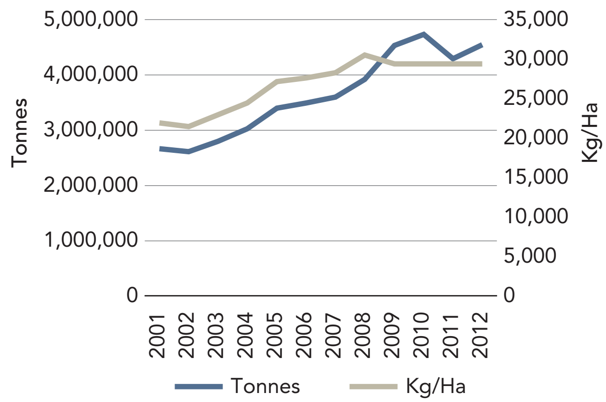
Sources: EPM 2001–2010 and ENSOMD 2012

FIGURE 1.6: Price of Rice Paddy (Producer Price in Communities) by Consumption Quintile



Sources: EPM 2005, 2010.

FIGURE 1.7: Rice Production and Yields



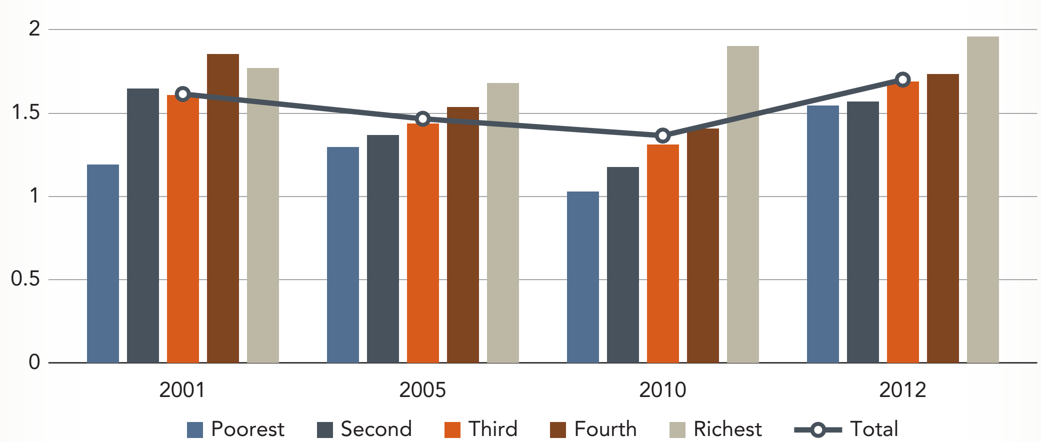
Source: FAO.

fertilizer (urea), which fell overall between 2005 and 2010, decreased by 18 percent.[11]

Although aggregate trends suggest that more land was brought under rice cultivation after 2008, the average productivity of this land fell. As shown in figure 1.7, rice yields flattened after the world food price spike of 2008, but production continued to increase through an expansion in the land under rice cultivation, in part due to continued high population growth.[12] Yet the average area cultivated per agricultural household continued its downward trend (figure 1.8) before reverting in 2012 to its 2001 level.[13] Over the decade agricultural productivity per worker fell (figure 1.9). This, combined with demographic trends, increased logging activities and weak enforcement, especially after the political crisis of 2009, may have exacerbated Madagascar's deforestation problem, already under way (figure 1.10).

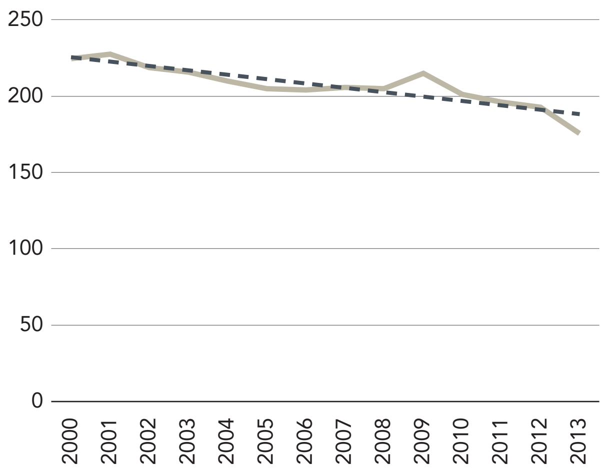
Households once again responded to circumstances in 2010 by shifting their employment patterns. As the terms of trade in agriculture deteriorated, a slightly lower percentage of households in the 3rd and 4th quintiles had heads primarily employed in agriculture, and a much

FIGURE 1.8: Area of Economically Exploited Land per Agricultural Household, by Year and Consumption Quintile (Hectares)



Sources: EPM and ENSOMD.

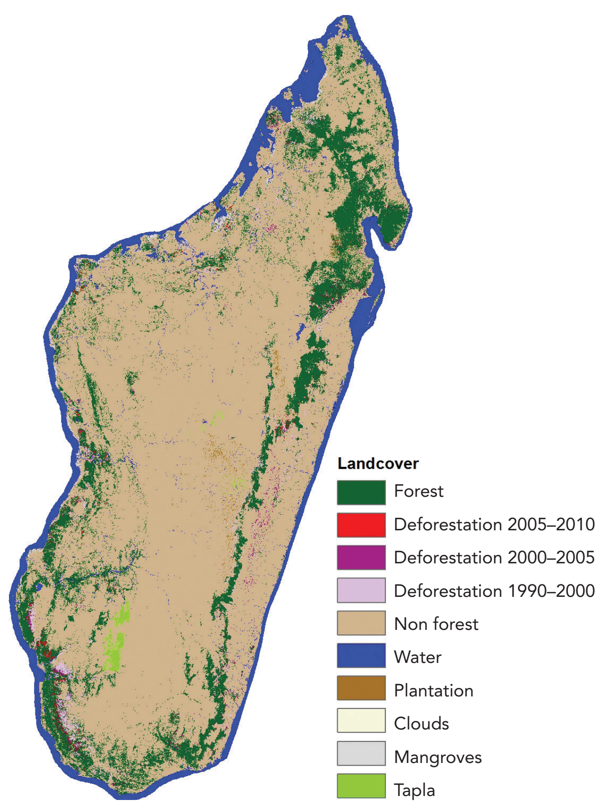
FIGURE 1.9: Agricultural Value Added Per Worker



Source: WDI.

lower percentage in the top quintile (figure 1.5). One observes an even more significant decline in agriculture as a sector of secondary employment between 2005 and 2010, in all quintiles. Among those with secondary employment, after agriculture, service sectors employed the greatest percentage of all household heads in 2010. And although services were an important source of secondary employment in all years, there was an especially dramatic increase in employment of household heads in this sector in 2010 (figure 1.11).

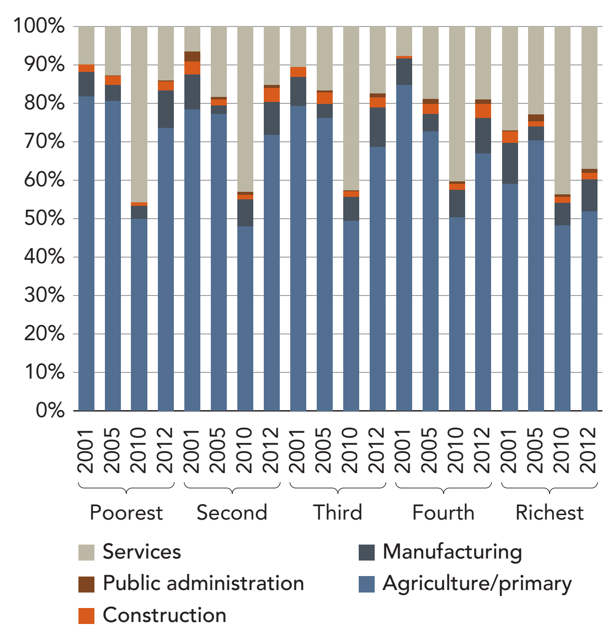
FIGURE 1.10: Forest Cover and Changes over Time



Trends in the main sectors of employment once again reversed between 2010 and 2012. In 2012 those in the bottom quintiles were more likely to be employed in agriculture than in 2010, but those in the top were more likely to be employed off-farm. The declining trend in manufacturing employment, which continued through 2010, began to reverse in 2012, when 10.2 percent of people in the top quintile had a household head primarily employed in manufacturing (versus only 2.2 percent in the bottom quintile). The entry into services in 2010 also partially reversed in 2012 (figure 1.11)

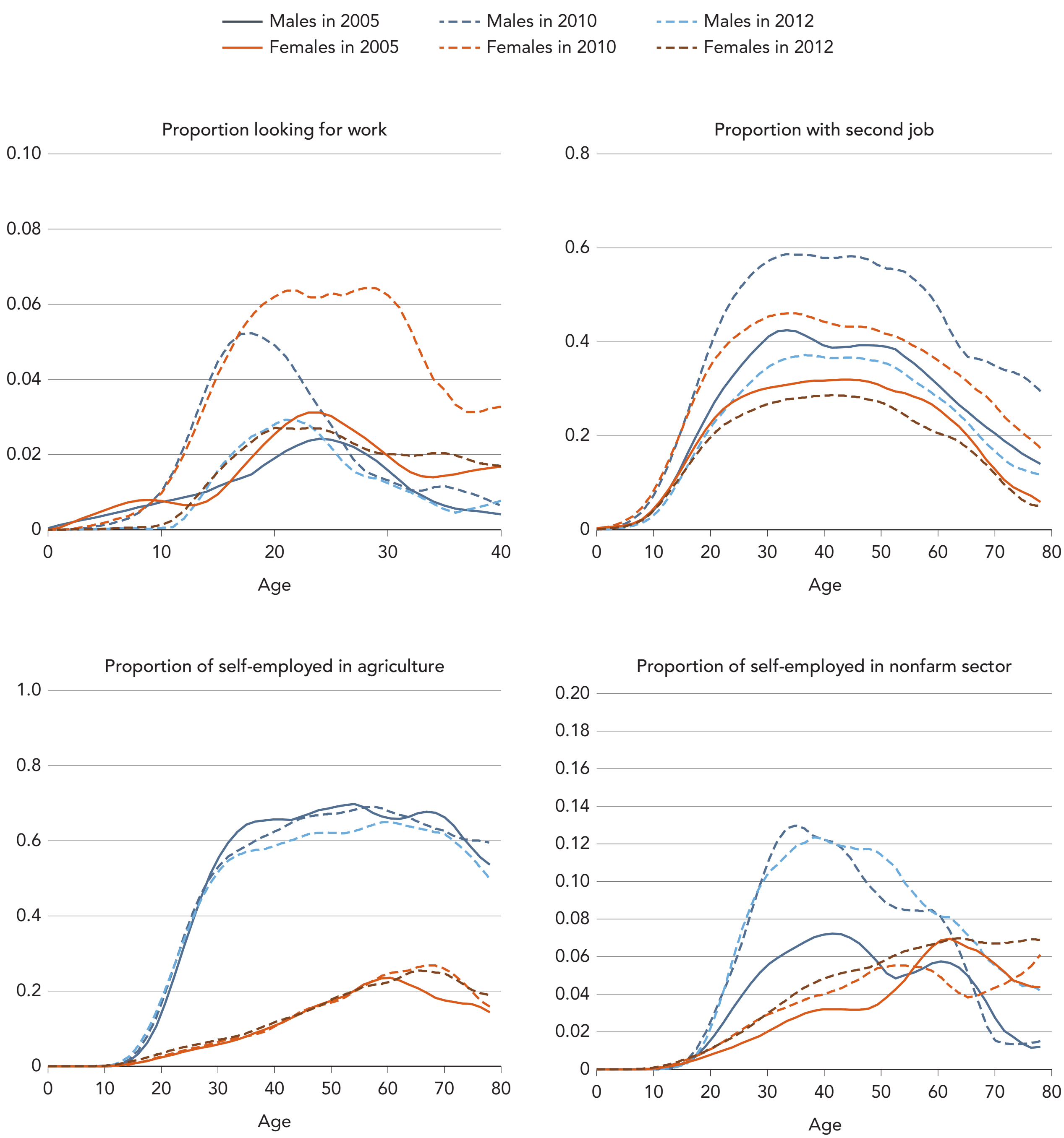
Households also responded to events by seeking a second job and self-employment off-farm. As shown in figure 1.12, as agricultural profitability declined, the proportion of both males and females looking for work increased significantly between 2005 and 2010, especially for females over the age of 10. In 2012, however, those looking for work fell again for both genders. At the same time, the proportion of both males and females with a second job increased from 2005 to 2010, and

FIGURE 1.11: Sector of Secondary Employment of Household Head (%)



then dropped in 2012 to levels below those observed in 2005. Women were much less likely to be self-employed in agriculture than men. Yet the percentage of both male and female workers self-employed outside of agriculture increased between 2005 and 2010, and remained higher in 2012 for most age ranges. The proportion of men self-employed in agriculture also declined in 2012 after staying relatively constant for all ages between 2005

FIGURE 1.12: Labor Market Outcomes, 2005, 2010, and 2012 (Kernel-Weighted Local Polynomial Smoothed Age-Outcome Profiles



Source: Calculated using EPM and ENSOMD.
Note: *y* axis represents polynomial smoothed proportion (from 0 to 1) of individuals engaging in labor market behavior noted.

and 2010. A shift out of agricultural work may signal improving opportunities in off-farm labor markets and small informal enterprise; yet overall wages have not kept pace with inflation (see figure 1.13) and, as discussed in Bi and Osborne (2016), employment in the smallest of such enterprises is typically less productive and remunerative than in other jobs. Thus, labor productivity remains too low to make a greater dent in the country's poverty rate. Moreover, women have had more difficulty securing employment off-farm, and the disparity in wages between females and males of prime working age increased in 2010 vis-à-vis 2005 (see disparities at age 40 in figure 1.13, as indicated by the arrow).

Community informant surveys are broadly consistent with the trends and indicators observed in agriculture. The EPM 2010 surveyed key informant community members in each of 623 communities and obtained the groups' list of the top five development problems in agriculture. A count of the frequency of responses for the top constraints, as well as inclusion in the top three constraints, is shown in figure 1.14. Although these data are based on perceptions rather than quantitative analysis, a couple of themes clearly emerge. First is the importance of problems in input markets. The most frequently ranked issue among the top three constraints is related to either the high cost or lack of access to inputs such as seeds and fertilizer.[14] This was also the fourth most frequently cited among communities' number one issues.

Many communities also cited the lack of adequate land area as a key obstacle. Some 22.5 percent of communities ranked this as the number one constraint—the most frequent among the top ranked constraints cited—and a lack of land was ranked third in frequency among households' top three constraints. This likely reflects the high price and low profitability of inputs, which would incentivize extensive over intensive agriculture. A tally of responses shows that insecurity of land tenure and conflict over land were mentioned in only a few communities.[15]

Following these top issues, many communities (34 percent) ranked the condition of irrigation infrastructure, then the condition of roads (25 percent) as among their top three problems for agricultural development. In addition, insecurity was a major issue. In 22 percent of communities theft of cattle and in 15 percent of communities theft of crops were listed among the top three problems, and combined they pose a greater issue than the condition of roads. In addition, distance to product markets was cited in 19 percent of communities.[16]

Notably, issues related to access to credit did not rank high on communities' lists of priority problems. As shown, relatively few communities cited either the distance to credit institutions, credit security requirements, or high interest rates as among the top three agricultural development problems. According to community

FIGURE 1.13: Wage Trends from 2005 to 2012

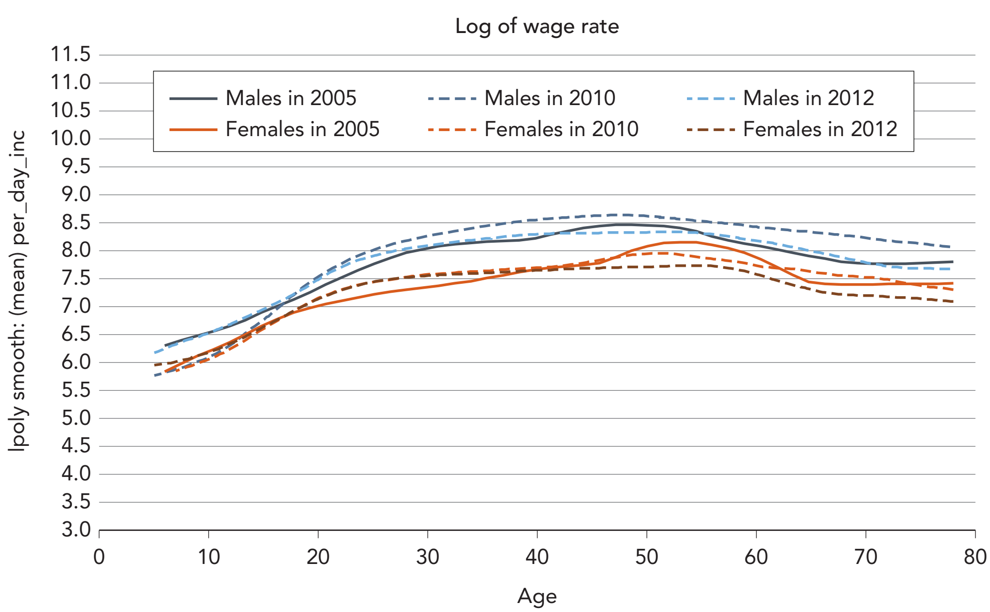
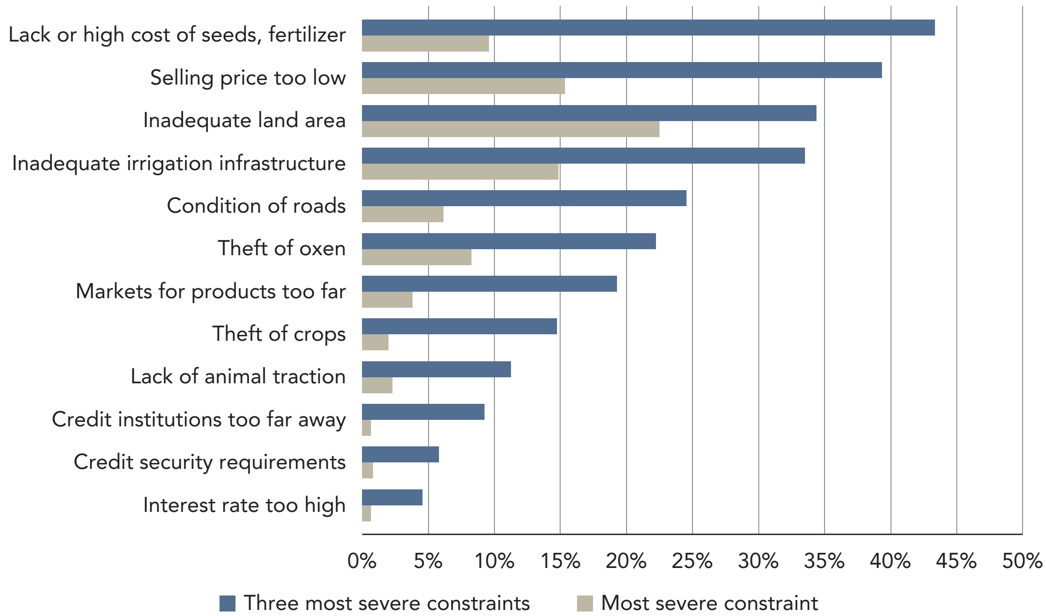


FIGURE 1.14: Frequency of Community Group Rankings of Constraints to Agriculture



Source: Calculated using EPM 2010.

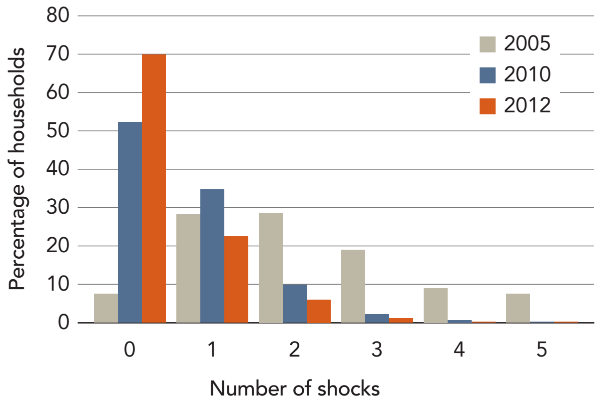
member perceptions, therefore, issues related to greater profitability—access to inputs and markets—rather than a lack of access to credit by producers are the key obstacles to agricultural development. Similarly, seasonality of labor demand, inadequate access to road and irrigation infrastructure, and weak market access appear among the country important constraints, based on studies conducted on Madagascar over the past 15 years.[17]

Risks and Vulnerability

Households in Madagascar are subject to an extreme amount of weather-related risk, which can push them deeper into poverty at any time, and these risks were most clearly manifested in 2010.[18] Although people tend to prefer a relatively even level of consumption over time even as income fluctuates, they cannot perfectly smooth short-term income fluctuations arising from weather, price, or temporary health shocks by borrowing, saving, and insuring against risk (see, for example, Friedman 1957).[19] Without further study, it is unclear to what extent informal risk-mitigation instruments are available in rural Madagascar, but most households report giving and/or receiving remittances. Even so, the available strategies are unlikely to adequately address spatially correlated risks such as cyclones or drought. Thus, consumption levels are likely to respond significantly to short-run shocks—as would inequality and poverty measures—without these changes necessarily being persistent (or indeed permanent).[20] For instance, when there is greater spatial variance of weather shocks in a given year, measured inequality in that period will appear higher, without this necessarily representing a permanent condition. At the same time, when such shocks are large and significant assets are lost, households will have difficulty recovering economically and may be forced to sacrifice long-run investments in education and health as part of their coping strategy.

Natural conditions that result in such huge intermittent losses combined with the absence of adequate mechanisms to shield against them not only have devastating short-run effects on consumption but also make it necessary to hedge risks in a way that persistently reduces incomes.[21] For example, farmers must avoid specialization and dependence on food markets and rather must operate in relative autarky: the percentage of crop production for the market is low, and one sees even urban households engaged in agriculture for their own production. Moreover, when combined with poorly performing input markets—the inability to access inputs at the right times and at advantageous prices—these issues reduce profitability substantially. The returns to using fertilizer, for example, in these circumstances can be nil (see Livingston et al. 2011).[22]

FIGURE 1.15: Number of Negative Shocks Reported (2005–12)

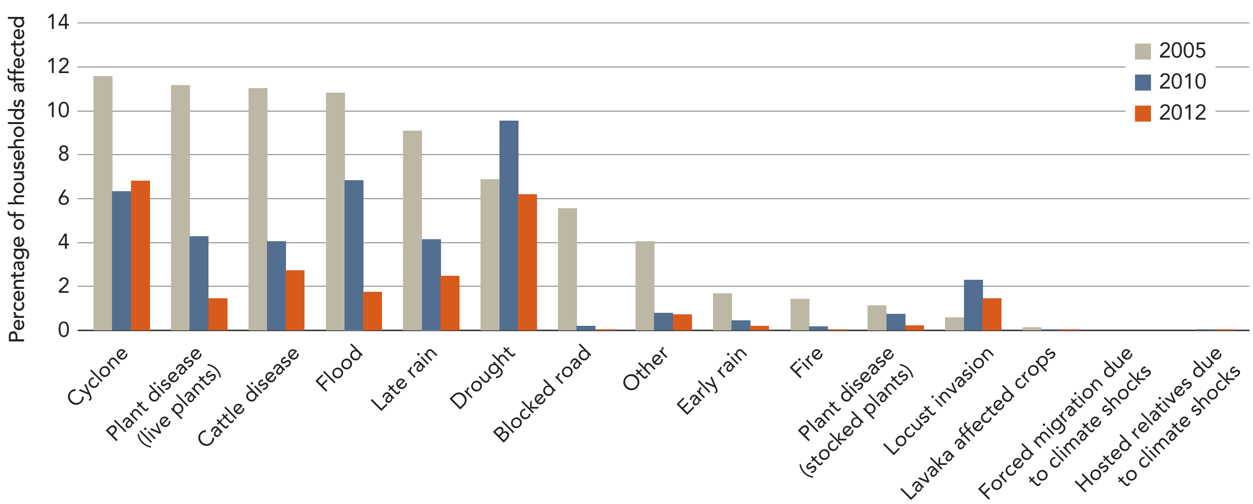


Source: EPM.

As shown in figure 1.15, the number of shocks—whether climatic, health, security, or economic shocks—that households experienced was lower in 2012 than in any of the other prior survey years. Apart from 2005, when a general price increase was the most frequently reported shock, climatic shocks are the most frequently reported type. As shown in figure 1.16, more households reported being affected by a cyclone, flood, or late rains in 2005 than in later years. Plant and animal disease also affects a significant percentage of households.

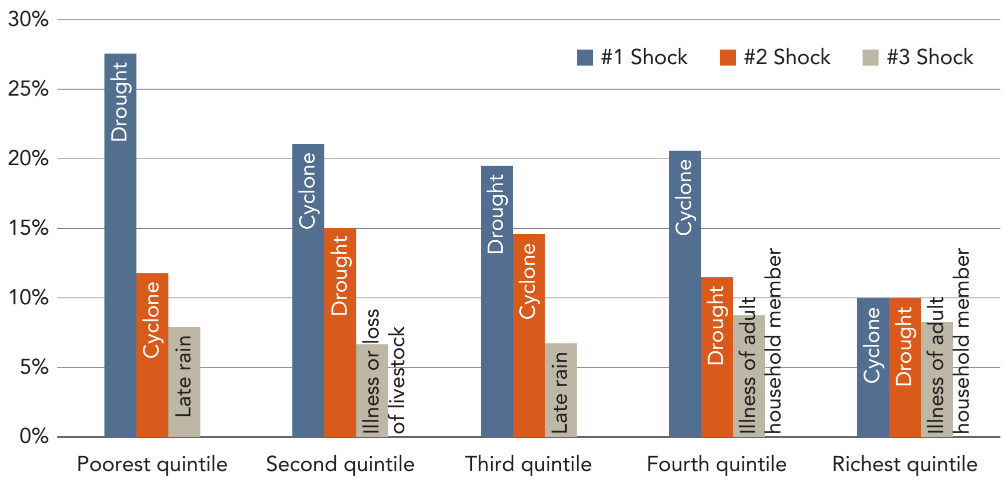
Moreover, although the type of adverse shock changes from year to year and affects different households, the frequency of adverse climatic shocks is generally correlated with poverty, as shown in figure 1.17. As shown in figure 1.18 2005 was also a bad year for health shocks relative to the subsequent survey years, as it was for security shocks (figure 1.19). Nonetheless, the costs of these shocks appear to have been greatest for the poorest households particularly in 2010, as shown by Thiebaud, Osborne and Belghith (2016). The full statistics on the frequency of shocks is reported in annex 1C.

FIGURE 1.16: Climatic, Natural, and Related Shocks



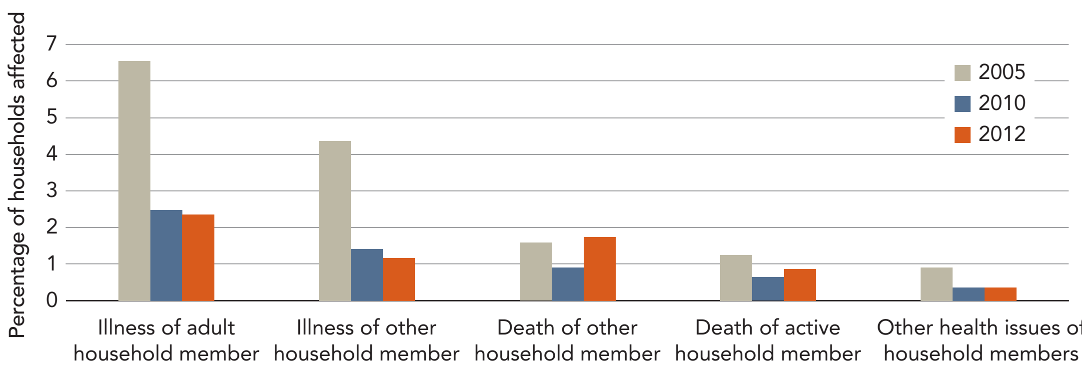
Source: EPM.

FIGURE 1.17: Percentage of Households Having Stated Shock (Top Three Reported Shocks, 2012)



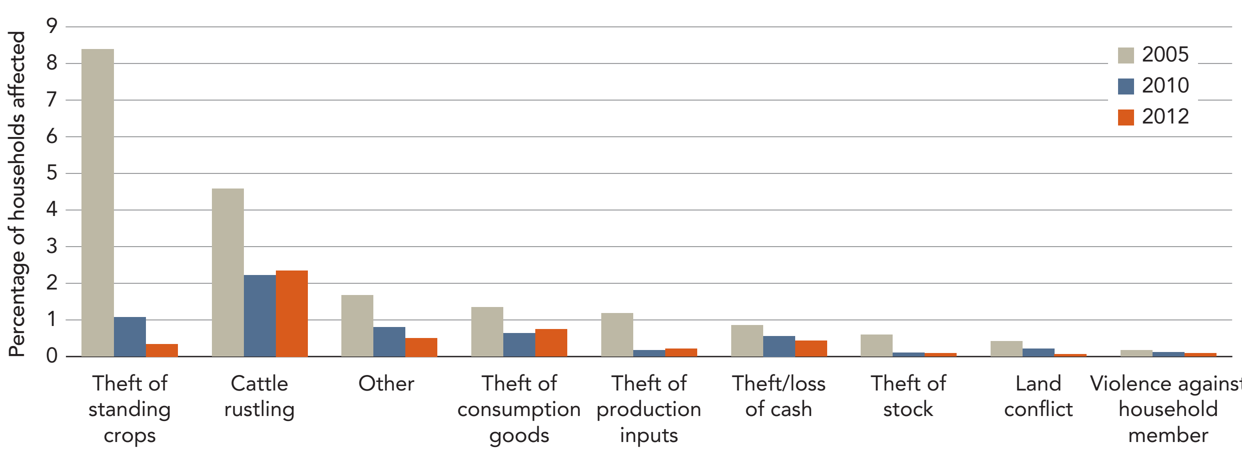
Source: ENSOMD 2012

FIGURE 1.18: Frequency of Health Shocks Reported



Source: EPM 2005, 2012, ENSOMD 2012.

FIGURE 1.19: Frequency of Reported Security Shocks



Source: EPM 2005, 2010, ENSOMD 2012.

Annex 1A. Sampling, Comparability between 2010 and 2012 Data and Weights Issues

The 2012 poverty statistics were calculated from Enquête Nationale sur les Objectifs Millenaire du Développement (ENSOMD), a household (HH) survey very similar in format to the country's previous Living Standards Measurement Surveys (LSMS), the Enquête Périodique auprès des Ménages (EPM). Because the objectives of the two surveys differed, so did the sampling strategy. The ENSOMD was designed to track outcomes related to the Millennium Development Goals, whereas EPM surveys are designed to capture a greater range of socioeconomic variables. For the EPM 2010, sample size and structure are similar to previous versions of the EPM, whereas for the ENSOMD 2012, sample size and sample structure are more similar to the Madagascar Demographic and Health Survey (DHS) survey. In particular, given the need for more detail on health outcomes and other indicators, ENSOMD did not include a community questionnaire. In order to capture DHS indicators such as mortality rates, the total HH sample size of the ENSOMD 2012 had to be very large.

TABLE 1A.1: Sampling Objectives for EPM and ENSOMD

Survey	Sampling objective
EPM 2010	To obtain a total HHs sample representative at the national level and at regions cross urban-rural. So, that sample is supposed to be representative for each of the 44 strata of the first stage sampling as there are 22 regions for Madagascar.
ENSOMD 2012	To draw a HH sample that is representative for the following domains: the national level, the capital, the urban and rural areas, and the 22 regions. The sample is not supposed to be representative for an urban-rural division within a region.

Sources: INSTAT, EPM 2010, and ENSOMD 2012 reports; Discussion with technical staffs of the surveys.

For the EPM 2010, the sample structure for urban-rural was driven generally by two criteria: geographic/regional population structure and the need to capture the diversity of socioeconomic life between rural and urban areas and within urban areas. Thus, the sample of EPM 2010 is almost equally distributed by urban-rural strata. For the ENSOMD 2012, the sample structure is mainly

TABLE 1A.2: Sampling Methodology for EPM 2010 and ENSOMD 2012

Rubric	EPM 2010	ENSOMD 2012	Observations
Sampling methodology	Two stages, EAs in first stage and HH in second stage	Two stages, EAs in first stage and HH in second stage	
Sampling frame for first stage	EAs of the census mapping of 2008	EAs of the census mapping of 2008	
Stratification in the first stage	Regions cross urban-rural areas	Regions cross urban-rural areas	The EPM 2010 retained the old definition of urban-rural, while the ENSOMD 2012 used the new definition.
Sampling method in the first stage	Probability proportional to size	Probability proportional to size	
Segmentation during the enumeration step[a]	No segmentation	Segmention is used for identified big EAs selected from the first stage	Segmentation is used only in the 2012 survey
Sampling method at second stage	Systematic sampling	Systematic sampling	

Sources: INSTAT, EPM 2010, and ENSOMD 2012 reports; Discussion with technical staffs of the surveys.
[a]For a selected EA in the first stage of sampling, segmentation denotes an action during the enumeration in the field; the field team divides the entire EA into two or more almost equal-size subdivisions. Afterward, the survey will be done in one subdivision randomly selected among all subdivisions of the EA. Segmentation is applied for EAs identified as large during the enumeration step by the survey team.

derived from the actual geographic partition of population. Only about 25 percent of the sample is drawn from the urban area in the 2012 survey, a level more similar to the actual population by area of residence. Moreover, the survey for 2012 is not representative of each region's urban and rural areas separately, as the 2010 survey is. Rather it is representative only by region and by rural and urban areas at the national level. Nonetheless, the core modules of the questionnaires, including consumption modules, were essentially identical across the two surveys.[23]

The two surveys use a two-stage sampling procedure wherein the first stage, a sample of Enumeration Areas (EAs), is randomly drawn from an EAs sampling frame and, in the second stage, a sample of HHs is drawn from a list of households obtained by an enumeration activity in each selected EA. In the first stage of sampling, the two surveys use the same sampling frame—the national list of EAs from the census mapping of 2008.[24] Although the sampling strategy differed, in principle as long as the sampling weights reflect the best available estimates of the population's structure, consumption aggregates and poverty numbers should be comparable at levels for which samples are representative. The sampling objectives for the two surveys are summarized in table 1A.1.

An additional complicating factor is that a new official definition of *urban* versus *rural* was applied beginning in 2012. The old definition of area of residence was the definition used since the last population census in 1993, whereas when producing the new database of EAs in 2008, Madagascar's National Institute of Statistics (INSTAT) used a new definition of urban-rural. For the most part, the reclassification of EAs led some urban areas in the old definition to be redefined as rural. Whereas the sampling for EPM 2010 still relied on the old definition, both old and new definitions could be captured in order to ease comparability. The ENSOMD 2012, however, used the new definition in the sampling frame. Table 1A.2 summarizes the sampling methodology used for the two surveys. Table 1A.3 show sthe resulting sample details, and 1A.4 shows the spatial pattern of the 2010 and 2012 samples.

TABLE1A.3: EPM 2010 and ENSOMD 2012 Sample Comparison

	EPM 2010	ENSOMD 2012
Initial sample size		
Sample of enumeration areas	623	615
Sample intake of HH by EA	20	32
Total HH sample	12,460	19,680
Final sample size		
Sample of EAs	623	609
Total HH sample	12,460	16,920

Source: INSTAT.

TABLE 1A.4: Partition of Sample of EAs by Region and Urban-Rural Area for Each Survey

	EPM 2010			ENSOMD 2012		
Region	Urban	Rural	Total	Urban	Rural	Total
Analamanga	30	24	54	50	25	75
Vakinankaratra	15	15	30	6	19	25
Itasy	12	13	25	3	22	25
Bongolava	12	13	25	4	21	25
Matsiatra Ambony	14	13	27	7	18	25
Amoron'i Mania	13	13	26	3	22	25
Vatovavy Fitovinany	14	14	28	3	23	26
Ihorombe	12	12	24	4	21	25
Atsimo Atsinanana	12	13	25	4	21	25
Atsinanana	19	14	33	7	19	26
Analanjirofo	13	14	27	4	22	26
Alaotra Mangoro	13	13	26	5	20	25

(continued)

TABLE 1A.4: Partition of Sample of EAs by Region and Urban-Rural Area for Each Survey *(continued)*

Boeny	17	14	31	8	17	25
Sofia	14	17	31	4	21	25
Betsiboka	12	12	24	3	22	25
Melaky	12	12	24	3	23	26
Atsimo Andrefana	14	18	32	6	20	26
Androy	12	12	24		26	26
Anosy	14	12	26	5	22	27
Menabe	14	12	26	5	21	26
DIANA	14	12	26	10	15	25
SAVA	14	15	29	5	20	25
Total	316	307	623	149	460	609
% of urban-rural	50.7%	49.3%	100.0%	24.5%	75.5%	100.0%

Sources: INSTAT, EPM 2010, and ENSOMD 2012 reports and databases; Discussion with technical staffs of the surveys.

TABLE 1A.5: Weight Construction Procedure and Components

Component	EPM 2010	ENSOMD 2012	Observations
Design weight for each EA	(Pop total in strata/pop in the EA) x Sample size of EAs in strata	(Pop total in strata/pop in the EA) x Sample size of EAs in strata	
Segmentation		1/(proportion of the segmentation)	Segmentation is not applied for 2010
Design weight for each HH at EA level	(Number of enumerated HH in the EA)/20	(Number of enumerated HH in the EA)/32	
Nonresponse adjustment for EAs		(Sample size of EAs in strata)/(Sample size of EAs surveyed in strata)	Nonresponse adjustment is not applied for 2010
Nonresponse adjustment for HHs	—	(Number of identified HHs as sample in strata)/(Number of HHs with completed interview in the strata)	Nonresponse adjustment is not applied for 2010
Post-stratification to take into account geographical structure of the population	Nothing done here	The structure of the population in the 2008 census mapping was used to calculate adjustment factor (Wi)	The only post-stratification adjustment done was on the geographical repartition of population and it was done only for the 2012 survey.
Others	The total of HH was adjusted for some EAs for which the total of HHs enumerated was too low or too high, compared to the size of the EA in the sample frame[a]		
Final HHs weight	Multiplication of each above component	Multiplication of each above component	

Source: INSTAT, EPM 2010, and ENSOMD 2012 weight construction templates files; Discussion with technical staffs of the surveys.
[a]For some EAs, the number of HHs effectively enumerated by the field work team was judged by the survey analyst team to be too low or too high given the initial size of these EAs as already reported the 2008 EAs database. To correct, the initial size in the EAs database was taken into consideration, but this correction was done for a just few number of EAs (45 EAs among the total of 623 EAs).

The resulting sample structure is quite different for the two surveys as shown in table 1A.4. Based on the templates files of weight construction of the two surveys, all components included in the final HH weight used for data analysis are described in table 1A.5 for each survey.

From table 1A.5, one can say that the main components of the weight adjustments represent corrections for each survey following the corresponding sampling method. Nonetheless, there was a significant difference between the two surveys in that a poststratification to adjust the regional structure of population was done for the ENSOMD 2012, whereas this was not done for the EPM 2010. This resulted in an implicit population structure which differed from the best available information on the geographic allocation of the population.

Because standards of living vary importantly by area of residence and by region, we estimate the population structure by area of residence and by region in order to check the consistency of the actual weight of each of the 2010 and 2012 surveys, taking the structure from the 2008 EAs database as the definitive reference. As the last effective population census was done in 1993, this database is the most recent and best estimate available of the geographic structure of the Malagasy population. Figure 1A.1 compares the structure of the population by area of residence (using the old definition) from the 2008 census mapping with results from previous EPM surveys.

When deciding on the assumed structure of the population for the purposes of the 2010 survey, INSTAT conducted a statistical test of differences between 1993 and 2008 and was not able to reject that the structure remained the same. Nonetheless, for both 2010 and 2012 surveys, INSTAT utilized the slightly changed structure obtained from the 2008 census mapping exercise. INSTAT has not, therefore, modelled trends in population changes or urban-rural migration since that time for the purpose of altering the assured structure. These assumptions may be updated after a new census is completed, and there are no clear indications that the rural-urban structure has altered appreciably since 2008. Therefore, the best approach appears to be to hold the structure constant in calculating sampling weights and the corresponding statistics from the 2010 and 2012 surveys.

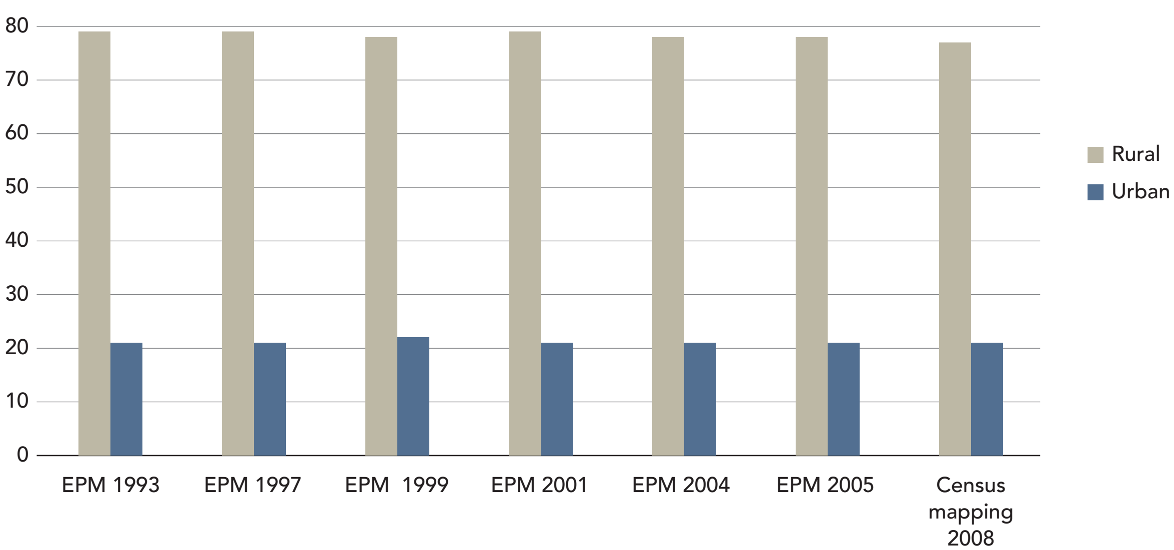
Table 1A.6 exhibits results obtained for the population's structure by urban-rural areas. To avoid confusion, results are shown separately for the old and the new definition of area of residence.

TABLE 1A.6: Structure of Population by Urban-Rural EPM 2010 and ENSOMD 2012

	Old definition			New definition		
Area	EA database	2010	2012	EA database	2010	2012
Urban	22.4	20.3	24.5	16.7	10.6	20.1
Rural	77.6	79.7	75.5	83.3	89.5	80.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

Sources: Calculated from EPM 2010, ENSOMD 2012, and census mapping of 2008 databases.

FIGURE 1A.1: Structure of Population by Urban-Rural from Previous EPM and the 2008 Census Mapping (Old Definition Rural-Urban)



Sources: INSTAT, EPM 1993, 1997, 1999, 2001, 2004, and 2005–2008 census mapping database of EAs.

The table reveals a major discrepancy in the percentage of rural and urban populations between 2010 and 2012, and in the case of the new definition of urban-rural, a discrepancy in 2010 between the percentage of rural and urban populations of approximately 6 percentage points. The partition of the population by region also shows some differences, as shown in table 1A.7. It appears clear from this that the structure of population provided by the EPM 2010 is problematic, while those provided by the ENSOMD 2012 seem reasonable, given that post-stratification adjustments were made for that survey. The importance of the discrepancies is exemplified by the proportion of population in the large region of Analamanga, which contains the capital city.

SUMMARY AND ADJUSTMENTS MADE

The main conclusion of the previous section is that weights applied in both the 2010 and 2012 surveys provide an inaccurate structure of population by urban-rural area of residence, whatever the definition used (old or new). In addition, the EPM 2010 actual weight does not provide a representative repartition of population by region. These issues need to be addressed as households' living standards vary by the geographical location of the household.

To address these issues, the World Bank poverty team has computed and applied new HH weights for both surveys. In addition to core design weights, the following adjustments were introduced:

1. A first post-stratification procedure for the two surveys for EA weight. In fact, the EA weight must reproduce the structure and the size of the sampling frame and it must be checked and corrected if not met.
2. A recomputed nonresponse adjustment at the EA level and at the HH level by EA.

TABLE 1A7: Structure of Population by Region, Census Map Reference versus EPM 2010 and ENSOMD 2012

Region	EA database	2010	2012
Analamanga	15.3	11.6	15.4
Vakinankaratra	8.3	8.3	8.3
Itasy	3.4	3.7	3.4
Bongolava	2.1	2.1	2.1
Matsiatra Ambony	5.5	6.0	5.5
Amoron'i Mania	3.3	3.4	3.3
Vatovavy Fitovinany	6.5	6.9	6.5
Ihorombe	1.4	1.2	1.4
Atsimo Atsinanana	4.1	4.4	4.1
Atsinanana	5.8	6.0	5.8
Analanjirofo	4.7	4.6	4.6
Alaotra Mangoro	4.7	4.6	4.7
Boeny	3.7	3.4	3.7
Sofia	5.7	5.6	5.7
Betsiboka	1.3	1.9	1.3
Melaky	1.3	1.4	1.3
Atsimo Andrefana	6.0	6.6	6.0
Androy	3.4	4.0	3.3
Anosy	3.1	3.1	3.0
Menabe	2.7	3.0	2.9
DIANA	3.2	2.8	3.3
SAVA	4.5	5.6	4.3
Total	100.0	100.0	100.0

Sources: Calculated from EPM 2010, ENSOMD 2012, and Census mapping of 2008 databases.

3. A post-stratification component to correct at the same time the structure of population by urban-rural and by region. This took account of the old definition of area of residence, the new definition of area of residence, and region.

Annex 1B. Poverty Estimation Methodological Notes

As it is typically the case in Sub-Saharan Africa, the available HH surveys in Madagascar use consumption as the key welfare measure to analyze poverty. This consumption aggregate comprises food consumption, including food produced by households themselves, as well as expenditures on a range of nonfood goods and services (for example, clothing, utilities, transportation, communication, health, education, housing-related expenditures, and imputed rent). However, the consumption aggregate does not include expenditures on larger consumer durable items (such as cars, TVs, computers, and so forth), nor does it include expenditures on ceremonies (marriage, funerals, and the like). To the extent that better-off households devote a larger proportion of their total consumption to durable goods, this omission creates certain biases and underestimates "true" consumption among wealthier families. This matters less for poverty analysis, where the focus lies on the bottom-end of the distribution, but it can have a significant impact on estimated inequality.

The HH surveys collect consumption data at the household level. For the purpose of poverty and welfare analysis, total HH consumption needs to be adjusted for differences in household size and composition, which imply different consumption expenditure levels to achieve the same utility. There is a "public" good aspect to some categories of consumption: for example, for housing and utilities, and different ages may require different nutritional intake. However, the approach followed here consists of computing consumption per capita, implicitly assuming that all members of the household require the same level of consumption, as this is the metric used by INSTAT as well as entities in many other SSA countries. Paasche price indices are used to adjust consumption per capita for differences in prices across geographic regions. The price indices are estimated using unit values from the surveys.

The poverty lines are based on the cost-of-basic-needs approach. The food poverty line is based on the cost of a food basket that delivers 2,133 calories per capita (given consumption patterns in a reference population) (see World Bank 2014). The basic needs poverty line adds an allowance for basic nonfood necessities to the food poverty line. The poverty lines have been reestimated for each survey year 2001, 2005, and 2010. This reestimation is done because there was a socioeconomic crisis that occurred between these years, which may affect the structure of consumption. Moreover, there is a rule of thumb according to which poverty lines need to be reestimated at least every five years. The poverty line for 2012 was estimated using the 2010 poverty line adjusted by the national consumer price index.

The basic needs headcount poverty rate (or, as used in the text, "poverty rate") measures the proportion of the population whose monthly (price-adjusted) total household consumption per capita is below the basic needs poverty line, and the extreme headcount poverty rate (used in the text as "extreme poverty rate") measures the proportion of the population whose monthly (price-adjusted) total household consumption per capita is below the food poverty line. The annual consumption poverty lines for each year covered in this report is as shown in the table 1B.1. Further technical details can be found in the Madagascar Poverty Assessment (World Bank 2014).

The national poverty line(s) reflect(s) Madagascar's specific costs of basic consumption needs, but they are difficult to compare with other countries' poverty thresholds.

TABLE 1B.1: Poverty Lines Used (Annual Consumption per Capita)

Year	2001	2005	2010	2012
Currency	MGA*	MGA	MGA	MGA
Food poverty line	734,320	227,085	294,690	341,840
Complete poverty line (nominal values)	963,554	289,169	381,791	442,877
Temporal deflator**	1	1.501	1.32	1.16

* 1 MGA (ariary) = 5 FMG (Malagasy franc, the former national currency replaced by the MGA from 2005 onward).
** Current survey compared to the previous survey year for all years.

To overcome this issue, the international poverty line of US$1.9 per capita per day (in 2011 PPP exchange rate) is often used to evaluate a country's poverty record vis-á-vis other developing countries or regions.

The poverty estimates for 2001, 2005, and 2010 in this paper differ from the poverty rates in the Madagascar Poverty Assessment due to the adjustment in the population weights as described above. The World Bank (2014) poverty figures in the report can be obtained exactly using the variables in the data and the old weights. The correction of the weight variable using the same post-stratification procedure described above has been applied to 2012 data.

Annex 1C. Detailed Data and Results Tables

TABLE 1C.1: Poverty Headcount and Distribution of the Poor by Region

	Poverty headcount rate					Distribution of the poor				
	2001	2005	2010	2012	Change	2001	2005	2010	2012	Change
Poverty line = Poverty line World Bank										
Urban	34.1	40.8	29.8	35.5	5.7	7.7	9.2	7.0	8.5	1.5
Rural	77.7	79.6	80.1	77.9	–2.2	92.3	90.8	93.0	91.5	–1.5
Region										
Analamanga		47.1	39.1	41.5	2.4		9.8	8.4	9.1	0.7
Vakinankaratra		83.3	77.6	87.7	10.1		9.6	8.8	9.9	1.1
Itasy		77.9	83.7	75.0	–8.6		3.7	3.9	3.4	–0.4
Bongolava		75.7	73.9	76.1	2.2		2.1	2.2	2.3	0.1
Matsiatra Ambony		84.5	81.0	71.9	–9.1		6.5	6.1	5.4	–0.7
Amoron'I Mania		86.1	85.9	81.7	–4.2		4.0	3.8	3.6	–0.2
Vatovavy Fitovinany		83.6	88.9	79.4	–9.5		7.3	8.1	7.4	–0.7
Ihorombe		84.4	79.1	76.6	–2.5		1.6	1.6	1.7	0.0
Atsimo Atsinanana		87.8	94.3	93.6	–0.7		4.8	5.5	5.7	0.2
Atsinanana		70.0	72.9	67.0	–5.9		5.7	5.8	5.4	–0.5
Analanjirofo		82.4	80.1	77.1	–3.0		5.4	5.3	5.1	–0.2
Alaotra Mangoro		66.6	72.3	62.8	–9.5		4.3	4.7	4.1	–0.6
Boeny		49.1	57.8	57.3	–0.5		2.4	3.0	3.1	0.1
Sofia		90.0	79.4	82.4	3.0		7.0	6.3	6.7	0.3
Betsiboka		76.9	81.9	78.9	–3.0		1.4	1.5	1.5	0.0
Melaky		81.1	79.0	81.6	2.6		1.4	1.5	1.6	0.1
Atsimo Andrefana		76.7	76.5	79.7	3.2		6.3	6.4	6.8	0.3
Androy		89.9	92.6	96.8	4.2		4.0	4.4	4.8	0.4
Anosy		76.0	78.5	85.7	7.2		3.2	3.3	3.7	0.3
Menabe		70.5	68.5	67.4	–1.1		2.5	2.6	2.7	0.0
Diana		51.8	46.2	36.4	–9.8		2.2	2.1	1.7	–0.4
Sava		76.0	71.2	71.9	0.7		4.7	4.4	4.5	0.0
Total	70.8	73.2	71.7	70.7	–0.9	100.0	100.0	100.0	100.0	0.0

(*continued*)

	Poverty headcount rate					Distribution of the poor				
	2001	2005	2010	2012	Change	2001	2005	2010	2012	Change
Poverty line = Food poverty line World Bank										
Urban	22.4	28.3	18.3	22.7	4.3	5.9	7.8	5.3	6.7	1.4
Rural	67.7	66.0	66.4	64.5	–1.8	94.1	92.2	94.7	93.3	–1.4
Region										
Analamanga		33.7	23.3	29.1	5.8		8.5	6.2	7.9	1.7
Vakinankaratra		68.7	61.2	78.0	16.7		9.7	8.5	10.8	2.3
Itasy		63.0	71.0	49.8	–21.1		3.6	4.0	2.8	–1.2
Bongolava		57.1	56.7	56.2	–0.5		2.0	2.1	2.1	0.0
Matsiatra Ambony		66.4	70.4	53.9	–16.5		6.3	6.5	4.9	–1.6
Amoron'I Mania		73.0	72.7	63.1	–9.5		4.1	4.0	3.4	–0.6
Vatovavy Fitovinany		72.3	76.2	66.6	–9.6		7.8	8.5	7.6	–0.9
Ihorombe		73.0	65.9	66.0	0.1		1.6	1.7	1.8	0.1
Atsimo Atsinanana		79.6	88.7	88.8	0.1		5.3	6.4	6.7	0.2
Atsinanana		59.5	61.1	53.4	–7.7		5.9	6.0	5.3	–0.7
Analanjirofo		72.5	68.9	60.5	–8.4		5.8	5.6	4.9	–0.6
Alaotra Mangoro		48.5	58.1	38.5	–19.6		3.9	4.6	3.1	–1.6
Boeny		37.1	40.6	46.2	5.6		2.2	2.6	3.1	0.5
Sofia		78.9	63.6	72.3	8.8		7.5	6.2	7.2	1.0
Betsiboka		58.6	69.1	57.3	–11.8		1.3	1.6	1.3	–0.3
Melaky		62.9	62.4	68.1	5.6		1.3	1.5	1.7	0.2
Atsimo Andrefana		66.1	65.3	72.8	7.5		6.7	6.7	7.6	0.9
Androy		81.0	84.9	92.1	7.2		4.4	5.0	5.6	0.6
Anosy		59.1	70.7	73.6	2.9		3.1	3.7	3.9	0.2
Menabe		51.8	51.6	52.3	0.7		2.3	2.4	2.6	0.1
DIANA		34.3	29.1	23.4	–5.7		1.8	1.6	1.3	–0.3
SAVA		63.4	58.5	56.7	–1.8		4.8	4.5	4.3	–0.1
Total	60.5	59.8	58.3	57.4	–0.9	100.0	100.0	100.0	100.0	0.0

Sensitivity of headcount poverty rate with respect to the choice of poverty line

	2001		2005		2010		2012	
	Poverty headcount rate	Change from actual (%)	Poverty headcount rate	Change from actual (%)	Poverty headcount rate	Change from actual (%)	Poverty headcount rate	Change from actual (%)
Poverty line = Poverty line World Bank								
Actual	70.8	0.0	73.2	0.0	71.7	0.0	70.7	0.0
+5%	73.0	3.1	75.5	3.1	73.6	2.8	73.0	3.2
+10%	74.8	5.7	77.4	5.8	75.8	5.7	74.9	5.8
+20%	77.7	9.8	80.9	10.5	79.3	10.6	78.4	10.9
–5%	68.9	–2.7	70.8	–3.3	69.4	–3.1	68.3	–3.4
–10%	67.1	–5.2	68.0	–7.2	66.9	–6.6	65.5	–7.4
–20%	62.4	–11.8	60.9	–16.9	60.4	–15.7	59.2	–16.3
Poverty line = Food poverty line World Bank								
Actual	60.5	0.0	59.8	0.0	58.3	0.0	57.4	0.0
+5%	62.4	3.2	62.9	5.2	61.2	5.0	59.9	4.4
+10%	64.4	6.5	65.5	9.6	63.5	8.9	62.5	8.8
+20%	67.4	11.4	70.3	17.6	68.4	17.4	67.1	16.9
–5%	58.3	–3.6	56.6	–5.4	55.3	–5.1	54.4	–5.2
–10%	55.1	–8.9	53.2	–11.1	51.7	–11.3	51.0	–11.2
–20%	49.0	–19.0	44.8	–25.1	43.8	–24.8	44.4	–22.7

TABLE 1C.2: Percent of Households Reporting Stated Shock (Ordered by Most Frequently Reported in 2005)

Climatic shocks	2005	2010	2012
Cyclone	11.59	6.33	6.82
Plant disease (live plants)	11.17	4.28	1.46
Cattle disease	11.04	4.05	2.71
Flood	10.84	6.84	1.76
Late rain	9.09	4.16	2.48
Drought	6.88	9.52	6.19
Blocked road	5.55	0.2	0.03
Other	4.04	0.78	0.72
Early rain	1.68	0.45	0.21
Fire	1.42	0.16	0.05
Plant disease (stocked plants)	1.14	0.74	0.23
Locust invasion	0.59	2.31	1.46
Lavaka-affected crops	0.14	0.01	0.01
Forced migration due to climate shocks		0	0
Hosted relatives due to climate shocks		0.02	0.01
Security shocks	**2005**	**2010**	**2012**
Theft of standing crops	8.4	1.08	0.35
Cattle rustling	4.59	2.23	2.36
Other	1.68	0.79	0.51
Theft of consumption goods	1.34	0.63	0.75
Theft of production inputs	1.19	0.17	0.21
Theft/loss of cash	0.86	0.56	0.44
Theft of stock	0.59	0.11	0.08
Land conflict	0.41	0.21	0.06
Violence against household member	0.17	0.12	0.09
Health shocks	**2005**	**2010**	**2012**
Illness of adult household member	6.54	2.47	2.35
Illness of other household member	4.37	1.41	1.16
Death of other household member	1.58	0.9	1.74
Death of active household member	1.25	0.64	0.87
Other health issues of household members	0.9	0.36	0.36
Economic shocks	**2005**	**2010**	**2012**
General consumer price increase	77.7	2.91	0.81
Increase in product prices	12.87	1.84	0.39
Increase in input prices	7.25	2.42	0.25
Death of person in community	2.35	0.53	0.09
Difficulty finding buyers of agricultural products	2.12	0.33	0.29
Other	1.82	0.55	0.19
Difficulty finding buyers of nonagricultural products	0.98	0.38	0.24
Loss of job of household member	0.76	0.85	0.28
Loss of animal used for traction		0.08	0.04
Farmgate prices too low		0.99	0.09
	2005	**2010**	**2012**
Other shocks	1.3	0.61	2.15

NOTES

1. The 2012 poverty statistics were calculated from *Enquête Nationale sur les Objectifs Millenaire du Développement* (ENSOMD), a household survey very similar in format to the country's previous Living Standards Measurement Surveys (LSMS), the *Enquêtes Auprès les Ménages* (EPM).
2. The 95 percent confidence interval was calculated for changes from 2000 to 2010 in the World Bank's extreme and absolute poverty headcount ratios, and this interval is fairly wide for the extreme poverty rate. This suggests an even wider confidence interval for the national poverty line, given that the margin of error increases as the poverty line approaches the mode of the distribution.
3. A variety of data treatment issues were addressed in the process of verifying poverty estimates for 2012, but we highlight the main factor here.
4. The assumed population structure is based upon a 1993 census, updated by a 2008 census mapping of households. However, the reliability of Madagascar's statistics is compromised by the lack of a more recent census.
5. PovcalNet 2012 data. This statement refers to countries with poverty data only.
6. Although PovcalNet does not use spatial price deflators, one can estimate the poverty rate using such deflators at the international poverty line, and one obtains a rate of 78.4 percent of the population in extreme poverty and 91.6 percent poor (living under US$3.10 2011 PPP) in 2012.
7. Although the survey instruments available do not allow us to update the geographic distribution of the population on a frequent basis, fluctuations in urban and rural poverty rates can be partially the result of migration of poor households to and from urban areas.
8. GDP growth estimates and poverty and consumption estimates are derived from different sources of data—the former from the national accounts of a country and the latter from household surveys—and very often estimated income and consumption diverge between these two sources.
9. Based on extreme poverty line (World Bank 2014).
10. The Gini coefficient is equal to the area between the Lorenz curve and the 45-degree line divided by the sum of this area and the area under this curve, and is expressed as $G = \frac{1}{\mu}\int_0^\infty F(y)(1-F(y))dy$, where μ is mean income/consumption and $F(x)$ is the distribution of income/consumption.
11. Staff calculations (bivariate regression) using EPM 2010.
12. Madagascar's high population growth rate, estimated at 2.78 percent (relative to a SSA mean of 2.71).
13. Since the underlying data is not a panel, one cannot conclude that poor households lost and then regained access to land over time. Average cultivated area in 2012 was 1.68 hectares versus 1.61 in 2001, but this is not a statistically significant difference.
14. This was computed as an aggregation of possible responses: lack of seeds, lack of improved seeds, lack of fertilizer, high cost of inputs, high cost of seed, and so forth.
15. The phrasing of the questionnaire referred specifically to limitations to land area, and did not ask directly about insecurity of tenure or conflicts over land. It is possible that respondents blurred the issues of access and tenure security.
16. In addition, other issues, including farmers' knowledge or support for introducing new technologies, weather or climatic issues, and soil fertility were mentioned, but not in sufficient frequency to be included in the top constraints communities mentioned.
17. Market integration is in turn important for improving producer prices and seasonal price smoothing. Other constraints relate to seasonality of labor inputs Moser and Barrett (2003) find that a promising system of rice intensification (SRI), while requiring low external inputs, is difficult for most farmers to practice because the method requires significant additional labor input at a time of the year when liquidity is low and labor effort is already high.
18. The World Bank conducted a comprehensive vulnerability assessment in 2012 using data through 2010 (World Bank 2012).
19. Some theories predict consumption growth with income growth. See, for example, Carroll (1997). Demographic issues and lifecycle saving are not included in this analysis, as the timeframe for analysis is relatively short, and reliable data needed to study this aspect of saving are not available.
20. Dollar, Kleineberg, and Kraay (2016), for example, find that inequality tends to show mean reversion in cross country data.
21. See, for example, Christiansen and Dercon (2011), Osborne (2006), and Zimmerman and Carter (2003), which underscore the importance of risk in farmers' decisions to utilize lower-effort, lower-return technologies.
22. In a recent trial, on-time fertilizer applications registered value/cost ratios of greater than two in eight of the 21 cases, compared to a ratio of zero among those who received fertilizer late, and according to common rule of thumb value/cost ratios of greater than two are needed for farmers to adopt fertilizer into their production systems (Livingston et al. 2011).
23. In addition, the sampling strategy differed; and as shown in table 1A.2, sampling was not done with replacement—so that it is possible that there is a greater problem of selection bias in the sample. In fact, selection bias is a potential problem in both surveys if households that were either not included or replaced were systematically different from those that were included.
24. The census mapping of 2008 was done in preparation of the national population census that was supposed to take place in 2009 but was not undertaken due to the 2009 crisis.

REFERENCES

Bi, C. and T. Osborne. 2016. "Transactions Costs, Poverty, and Low Productivity Traps: Evidence from Madagascar's Informal Micro-Enterprise Sector." Washington, DC: World Bank.

Carroll, C. 1997. "Buffer Stock Saving and the Life Cycle-Permanent Income Hypothesis." *Quarterly Journal of Economics* 112 (1).

Christiansen, L. and S. Dercon. 2011 Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics* 96: 159–173.

David-Benz, Hélène CIRAD. 2011. "A Madagascar: les prix du riz flambent, sans rapport avec le marché international." Paris: Cirad, UMR Moisa.

Deaton, A. 1992. *Understanding Consumption*. Clarendon Lectures in Economics. Oxford, UK: Oxford University Press.

Dollar, D., K. Kleineberg, and A. Kraay 2016. "Growth is still good for the poor. *European Economic Review* 81: 68–85.

Friedman, M. 1957. *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.

INSTAT. 2013. *Enquête nationale sur l'emploi et le secteur informel—ENEMPSI 2012*. Etude nationale. Vice Primature chargée de l'économie et de l'industrie.

INSTAT. 2014. *Enquête national sur le suivi des objectifs du millénaire pour le développement à Madagascar*. Etude nationale. Vice Primature chargée de l'économie et de l'industrie.

Livingston, G., S. Schonberger, and S. Delaney (2011). "Sub-Saharan Africa: The State of Smallholders in agriculture." IFAD. Conference on New Directions for Smallholder Agriculture, 2011.

Minten, B. 1999. "Infrastructure, Market Access, and Agricultural Prices: Evidence from Madagascar," Markets and Structural Studies Division (MSSD) Discussion Paper 26, IFPRI, Washington, DC.

Osborne, Theresa. 2006. "Credit and Risk in Rural Developing Economies. *Journal of Economic Dynamics & Control* 30: 541–68.

Thiebaud, A., T. Osborne and N. Belghith. 2016. "Isolation, Crisis, and Vulnerability: A Decomposition Analysis of Inequality and Deepening Poverty in Madagascar (2005–2010)." In this volume.

World Bank. 2012. *Madagascar Three Years into the Crisis: An Assessment of Vulnerability and Social Policies and Prospects for the Future*. Washington, DC: World Bank AFR Social Protection Unit.

World Bank. 2014. *Face of Poverty in Madagascar: Poverty, Gender, and Inequality Assessment*. Washington, DC: World Bank, PREM Africa.

Zimmerman, F. J., and M. R. Carter. 2003. "Asset Smoothing, Consumption Smoothing, and the Reproduction of Inequality under Risk and Subsistence Constraints." *Journal of Development Economics* 71 (2): 233–60.

CHAPTER 2

Isolation, Crisis, and Vulnerability: A Decomposition Analysis of Inequality and Deepening Poverty in Madagascar (2005–2010)

Alessia Thiebaud
Theresa Osborne
Nadia Belhaj Hassine Belghith
June 2016
The authors acknowledge the significant contributions to this paper by Patrick Leon Randriankolona.

Summary

Between 2005 and 2010, Madagascar experienced a moderate decrease in its headcount poverty rate. However, over the same period, the poorest of the poor fell deeper into poverty, particularly in rural areas, and inequality increased. Using an unconditional quantile regression method proposed by Firpo, Fortin, and Lemieux (2009), differences in consumption between groups of interest (urban and rural households) and changes in consumption over time (between 2005 and 2010) are decomposed to identify the main drivers of deepening poverty and increasing inequality, particularly in rural settings. Urban-rural inequalities in 2010 are mostly explained by a disparity in household endowments, which include some household assets, characteristics, shocks, and community-level variables. Differences in such endowments explain 78 percent of the total consumption difference between the poorest quintiles in urban versus rural areas, while differences in returns explain the remaining 22 percent, but the role of returns increases and of endowments diminishes for the higher consumption quintiles. Among households in the bottom quintile, over three-fourths of the difference in consumption levels between urban and rural households were attributable to differences in household size and composition,[1] human capital, climate shocks, and distances to food markets. Toward the upper end of the distribution, the role of returns becomes more prominent. Among the urban and rural households in the top quintile, about half of the consumption gap (49 percent) is explained by differences in endowments, and the other half by differences in returns (51 percent). The key structural correlates with consumption "explaining" disparities between rural and urban areas are remoteness from urban areas and the level of education of the household head. While more investments in transport connectivity and education in rural areas would have a positive effect on consumption and reduce urban-rural inequality, to fully realize the potential returns to these investments would require greater opportunities for urban migration and employment, in addition to economic integration with urban areas.

In addition, we decompose changes in consumption between 2005 and 2010 by quintile. We find that the increased severity of weather shocks, which disproportionately affected rural households and those in the poorest quintiles in 2010, is identified as a key driver of the observed changes. Decreasing returns to cultivated land and to being located in rural areas are also identified as fundamental drivers. We find that households in the poorest quintiles experienced the largest consumption losses between 2005 and 2010. Losses were particularly large for the rural poorest and were explained primarily by an increased severity of climate shocks and

by falling returns to agriculture, the latter of which is associated with a deterioration in the producer price relative to input costs and deteriorating transport conditions. In particular, climate shocks explain a –5.3 percent average change in consumption over the period among households in the poorest quintile (–7.0 percent in rural areas). Decreasing returns to agriculture explain a consumption change of –5.7 percent for households in the poorest quintile (–6.4 percent in rural areas). Thus, these two factors overexplain the actual change. As with the rural-urban analysis, the issues of remoteness and difficulties accessing markets emerge as a key explanation for the decline in rural incomes between the two years. As transport conditions deteriorated and rice policies acted to suppress increases in rice prices, the terms of trade in agriculture plummeted.

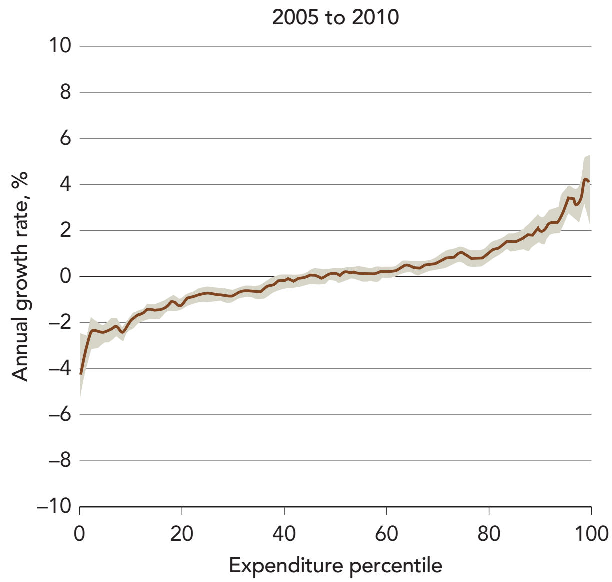
Offsetting these adverse effects on agriculture was a large increase in consumption unrelated to assets (except gender) for male-headed households relative to female headed ones, which are associated with a 13.8 percent increase in consumption for households in the bottom quintile (18.1 percent in rural areas). The net effect on consumption of households in the bottom quintile was a –3.1 percent change between 2005 and 2010 at the national level, and a –6.0 percent change in rural areas. We provide suggestive evidence that males were able to shift secondary work effort into services and other activities, whereas females faced more obstacles in doing so.

Introduction

Between 2005 and 2010, despite a modest decrease in Madagascar's national headcount poverty rate (from 73.2 percent in 2005 to 71.7 percent in 2010), inequality increased (see Belghith, Osborne, and Randriankolona 2016). The Gini coefficient rose from 38.9 to 42.7 and overall the incidence of growth was not favorable to the poor (figure 2.1). On a provincial level, poverty increased in 12 out of 22 provinces in Madagascar.

Moreover, there is an important rural-urban dimension to both persistent inequality and changes in consumption patterns over this period. The poverty gap in rural areas increased whereas that in urban areas decreased (see table 2.1). And as is typically the case in poor countries, poverty rates tend to be significantly higher in more rural provinces of Madagascar (see figure 2.2). Moreover, over the period of our study, we find that

FIGURE 2.1: Incidence of Consumption Growth (Total)



Source: Calculated using *Enquêtes Auprès les Ménages* (EPM) 2005, 2010.

it was the most rural provinces where average poverty rates increased the most. Whereas for the country as a whole the increase in inequality was mostly driven by higher consumption in the top quintile, in rural areas the increase in inequality was mostly due to a deterioration in consumption for the poorest households (figure 2.3).

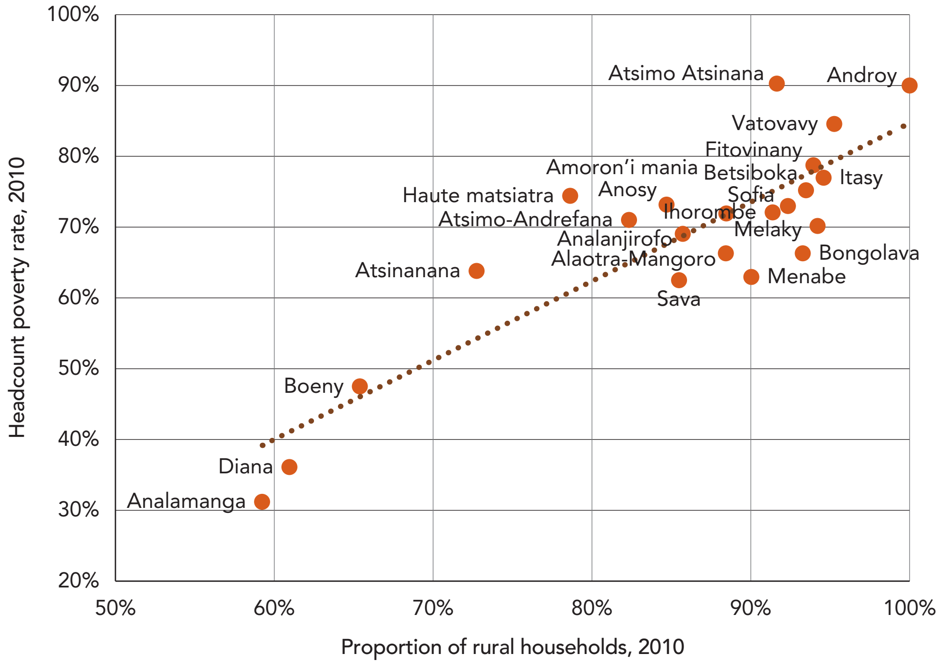
The objective of this paper is to provide a deeper empirical understanding of why consumption increased for some groups and not for others between 2005 and 2010, including why the largest consumption decline over the period occurred at the bottom of the distribution. Using recentered influence function (RIF) analysis, we uncover the main drivers of the increase in inequality—and changes in consumption levels—both over time and between urban and rural populations, for each quintile

TABLE 2.1: Trends in the Poverty Gap (Mean Percentage Shortfall of Consumption Relative to Poverty Line)

	2005	2010	Change
Urban	13.6	8.9	–4.7
Rural	34.8	36.7	+1.9
Total	31.3	32.0	+0.7

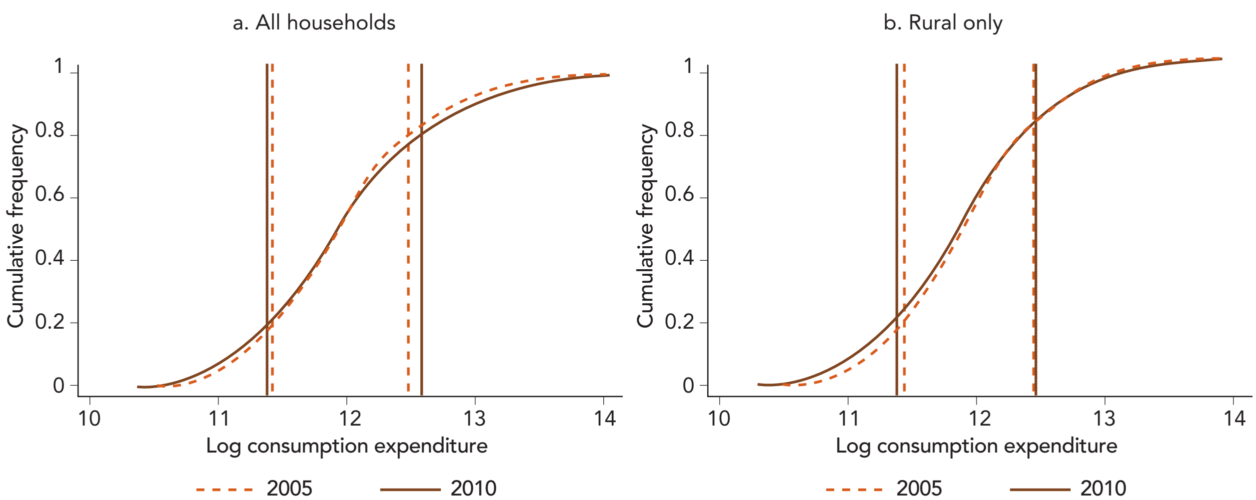
Source: Belghith, Randriankolona, and Osborne 2016.

FIGURE 2.2: Proportion of Rural Households and Headcount Poverty Rate (by Province)



Source: EPM 2005, 2010.

FIGURE 2.3: Cumulative Density Function of Log Consumption Expenditure



Source: EPM 2005, 2010.
Note: Cutoff points for bottom and top quintiles are indicated with vertical lines.

of the distribution. Because we utilize repeated cross-sectional data, the households falling into a given quintile will have shifted over time, and we cannot trace the persistence for given households of consumption or the effects of any influence variables on consumption. Rather, statements with respect to quintiles in different years relate to the respective quintiles for that year only.

We find that households in the poorest quintiles experienced the largest consumption losses between 2005 and

2010. Losses were particularly large for the rural poorest and were explained primarily by an increased severity of climate shocks and by falling returns to agriculture. In particular, climate shocks explain a –5.3 percent average change in consumption over the period among households in the poorest quintile (–7.0 percent in rural areas). Decreasing returns to agriculture explain a consumption change of –5.7 percent for households in the poorest quintile (–6.4 percent in rural areas). Thus, these two factors overexplain the actual change. Offsetting these was a large increase in consumption unrelated to assets (except gender) for male-headed households relative to female-headed ones, which are associated with a 13.8 percent increase in consumption for households in the bottom quintile (18.1 percent in rural areas). The net effect on consumption of households in the bottom quintile was a –3.1 percent change between 2005 and 2010 at the national level, and a change of –6.0 percent in rural areas. We suggest that males were able to shift secondary work effort into services and other activities, whereas females faced more obstacles in doing so.

Having identified the observed drivers of poverty and inequality, we attempt to relate them to the broader context, events, and policies. We find that a deep urban-rural divide continues to exist in Madagascar and is explained for the most part by differences in household endowments and characteristics, such as education level, distance to markets, and exposure to climate shocks. We find that sharp decreases in returns in rural areas and to cultivated land between 2005 and 2010, together with devastating effects of climate shocks, account for the majority of the drop in consumption experienced by the poorest households. We relate the decline in returns to rural areas and cultivated land to a context of low transmission of international food prices to poor Malagasy producers, increased transport costs, rising agricultural input costs, and deteriorating access to markets. We also highlight the role of climate shocks, which were more severe in 2010 than in 2005 and disproportionately affected the rural poor, contributing to significantly eroding their consumption levels. We identify increased participation in informal activities, particularly those pursued disproportionately by male entrepreneurs, as a primary means the poorest households used to offset these effects and avoid falling even deeper into poverty.

Overview of Poverty in Madagascar

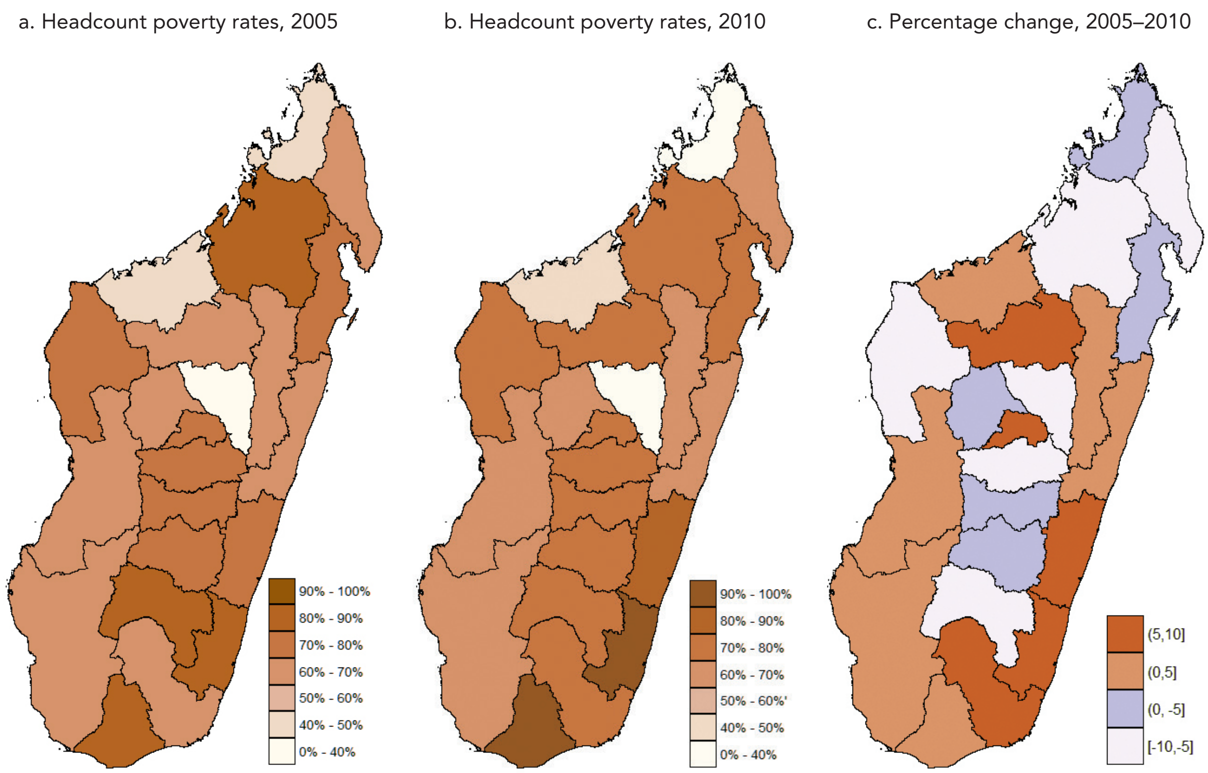
Poverty rates in Madagascar remain exceedingly high, particularly in rural areas, and progress toward poverty reduction has been slow. According to internationally comparable estimates, Madagascar's poverty rates are the highest in the world (Belghith, Osborne, and Randriankolona 2016). While events in urban areas are an important factor determining the headcount poverty rate, the protracted lack of progress in reducing extreme poverty in Madagascar is largely due to a failure to improve the lives of the rural poor, a vast majority of whom work in agriculture or the informal sector (usually both) (World Bank, 2015).

Between 2005 and 2010, despite a slight decrease in the overall poverty rate, the poorest fell deeper into poverty and inequality between the bottom and the top quintiles increased. Over the period of interest for this analysis, a modest decrease in the national headcount poverty rate was observed. The headcount poverty rate fell from 73.2 percent in 2005 to 71.7 percent in 2010. However, at the same time, inequality increased in Madagascar (see Belghith, Osborne, and Randriankolona 2016). The Gini coefficient rose from 38.9 to 42.7 and overall the incidence of growth was not favorable to the poor. On a provincial level, poverty increased in 12 out of 22 provinces in Madagascar (figure 2.4)

Poverty decreased in urban areas but increased in rural areas. The urban poverty rate fell from 40.8 to 29.8, and the poverty gap fell by 4.7 percentage points, from 13.6 percent in 2005 to 8.9 percent in 2010. On the other hand, rural poverty increased by from 79.6 to 80.1 while the poverty gap increased 1.9 percentage points (rising from 34.9 percent in 2005 to 36.7 percent in 2010 (Belghith, Osborne, and Randriankolona 2016). Poverty increased the most in the most rural provinces (figure 2.5).

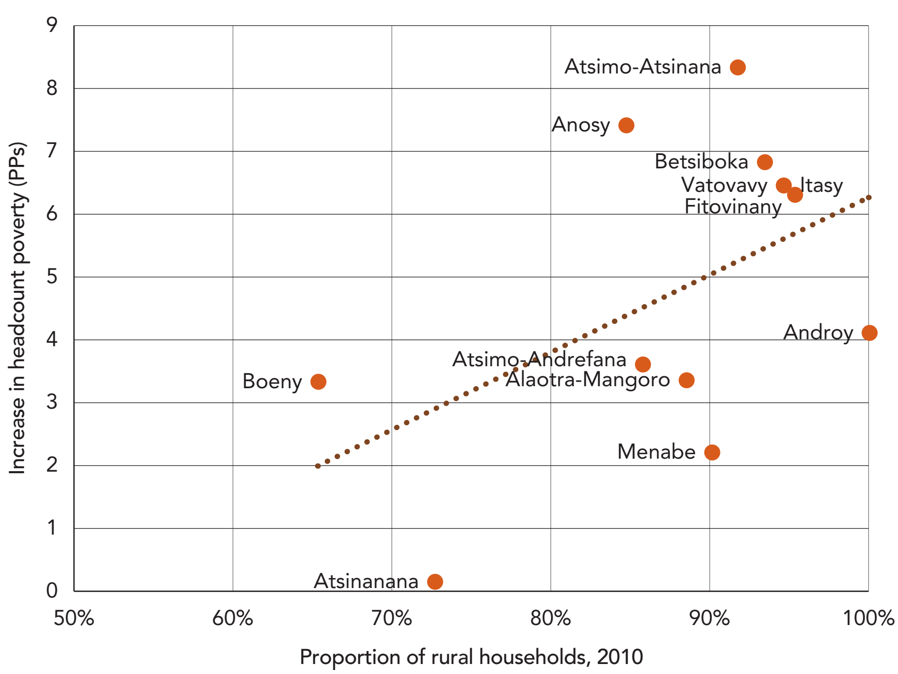
The rural poor and urban poor exhibit significant differences in household composition. As is generally the case in developing countries, the rural poor of Madagascar tend to live in households with more members, and with a higher proportion of children. They also have slightly younger household heads than their urban counterparts, on average. Poor rural households are also more likely to be headed by a male, and the household head is more likely to

FIGURE 2.4: Headcount Poverty Rates and Percentage Point Change (between 2005 and 2010, by Province)



Source: EPM 2005, 2010.

FIGURE 2.5: Proportion of Rural Households and Poverty Increase (2005–2010)



Source: EPM 2005, 2010. PP = percentage points.

TABLE 2.2: Summary Statistics for Urban and Rural Households (2010)

	Household size (Average number of members)		Age structure (Average members under 14)		Gender of head (Households with male head)		Age of head (Average years)		Marital status of head (Household heads with spouse)	
Quintile	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Bottom	5.6	6.2	46%	54%	72%	78%	42	41	72%	76%
Second	4.7	5.6	37%	49%	78%	82%	43	41	75%	80%
Middle	4.0	4.9	33%	44%	79%	82%	41	42	72%	79%
Fourth	3.6	4.3	25%	38%	76%	84%	41	42	69%	79%
Top	2.9	3.4	14%	24%	75%	80%	43	43	59%	70%
	Education level (Avg highest level completed by head/spouse, 1–4)		Health shocks (Households that had 1+ health shocks)		Climate shocks (Households that had 1+ climate shocks)		Distance to market (Households 1+ hours away from food market)		Security level (Avg security: 1 = very poor, 4 = very good)	
Quintile	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Bottom	1.9	1.4	9%	7%	23%	55%	18%	62%	3.1	3.1
Second	2.4	1.6	9%	4%	13%	41%	9%	53%	2.8	3.0
Middle	2.5	1.7	8%	4%	7%	37%	8%	51%	2.7	3.0
Fourth	2.9	1.8	6%	5%	7%	35%	7%	47%	2.7	3.0
Top	3.1	2.2	4%	6%	5%	27%	4%	33%	2.7	2.9

Source: EPM 2010.

have a spouse. Similar urban-rural differences are observed among richer households as well, as shown in table 2.2.

Poor rural and urban households also differ in the level of human capital and connectivity. Household heads in rural areas tend to be less educated than their urban counterparts, and this is observed for all quintiles. Rural households are also considerably more isolated than urban ones: while only 18 percent of urban households in the poorest quintile live one hour or more away from the food market, 62 percent of rural households in the poorest quintile do (table 2.2).

Poor rural and urban households have similar levels of health shocks and security, but poor rural households are significantly more affected by climate shocks. In 2010, 9 percent of urban households in the poorest quintile and 7 percent of rural households in the poorest quintile were affected by at least one health shock. Also, rural and urban households in the bottom quintile reported the same average level of security (a 3.1 score on a scale from 1 to 4, with 1 corresponding to "very poor" security and 4 corresponding to "very good" security). However, a sharp difference was observed in terms of climate shocks: while only 23 percent of urban households in the poorest quintile were affected by a climate shock, about 55 percent of their rural counterparts were, as shown in table 2.2.

Methodology

To uncover the proximate determinants of Madagascar's urban-rural inequality and changes in consumption between 2005 and 2010, we utilize the unconditional quantile regression method (based on the approach developed by Firpo, Fortin, and Lemieux 2009). This method can be used to identify the determinants of disparities in consumption expenditure both between socioeconomic groups and over time, and can be applied to each quantile of the distribution. This allows one to "explain" the distribution of consumption expenditure by a set of factors observed in household- and community-level data that vary systematically with socioeconomic status or that have varied over time.[2] The gap between groups or over time is decomposed into two parts: one due to group differences in the magnitudes of the variables associated with consumption levels ("determinants") and another due to group differences in the effects ("returns") of these determinants. This method allows us to identify the contributions of: (1) differences in household and community characteristics (endowment effects) and (2) disparities in returns to these characteristics (returns effect) to differences in consumption between groups and to changes in consumption over time at different quantiles. A more detailed discussion of the methodology is presented in annex 2B.

We carry out three separate decompositions. First, we decompose the differences in consumption between urban and rural households in 2010. Second, we decompose changes in consumption between 2005 and 2010 for all households. Third, we repeat this latter decomposition for rural households only, given the disproportionate deterioration in their living standards.

DATA AND VARIABLES

The Enquête Périodique auprès des Ménages (EPM), collected in 2005 and 2010, is used in this analysis. The EPM is a nationally representative household-level survey conducted by Madagascar's National Institute of Statistics (INSTAT). It provides extensive information on the demographic structure, education, health, employment, access to infrastructure, and consumption patterns of Malagasy households in both urban and rural settings. The EPM has been collected in 1993, 1997, 1999, 2001, 2002, 2004, 2005, and 2010. A national survey to monitor progress against the Millennium Development Goals (ENSOMD, in French) was also conducted in 2012 and is similar in its approach to collecting consumption data (see, for example Belghith, Osborne, and Randriankolona 2016). However, it was not suitable for this analysis as it lacks several key community-level variables.

In this exercise, household "endowments" are broadly defined. They include (1) household characteristics, such as household size, proportion of children in the household, gender of the household head, age of the household head, and marital status of the household head (whether in a couple or not); (2) human capital, as measured by education level of household head or spouse (whichever is higher); (3) shocks such as weather shocks and health shocks (dummy variables which are equal to one if the household has experienced at least one shock); (4) access to productive assets, including availability of electricity in the community (measured as the proportion of households in the community which have electricity, excluding the household itself) and availability of means of transportation; and (5) location, such as urban or rural setting, distance to the closest food market, security level (self-perception measure, on a scale from 1 to 5), land area under cultivation, and regional effects.

The choice of variables used in the unconditional quantile regression was made with the objective of mitigating concerns over simultaneity bias while attempting to explain as much of the differences in consumption as possible. Since the underlying model of consumption relies on permanent and temporary influences on households' real income and we seek to make causal (policy-relevant) inferences, particular attention was given to excluding variables that are less likely to be exogenous or predetermined. In particular, we wish to exclude variables that could themselves be affected by differences in returns (for example, sector of employment) or unobserved heterogeneity in ability or wealth (for example, ownership of assets that are primarily for consumption purposes), as these would bias the coefficients and make it impossible to infer causal "effects." For productive assets that are also likely correlated with wealth (for example, the specific type of transportation asset, household use of electricity), new variables were created to mitigate this problem. For example, a variable on the proportion of households within the community (excluding the household in question) that have electricity was preferred to a variable on the availability of electricity within each individual household. Cellphone use was excluded. Although having improved communication technologies can increase household incomes through a variety of channels, because more well-off households are more likely to adopt cell phones (and expenditures on utilities are likely easier to capture than other household expenditures), we suspect that the effect of unobserved wealth on household cellphone ownership would introduce bias on all coefficients. Similarly, the "effect" of having a car (versus other forms of transportation) would likely capture the effect of unobserved household wealth on consumption, in addition to the income gains possible from owning a car. We therefore include an indicator variable for whether or not the household owns any transportation asset, rather than variables to differentiate which type of asset this is. We acknowledge, however, that arguably all variables could be econometrically endogenous; previous educational, migration, and fertility decisions may be related to unobserved heterogeneity of the household. The exact specification used for each decomposition is shown in table 2.3. For all decompositions, full results are presented in the annexes.

TABLE 2.3: Decomposition Specifications

Variables	(1) Urban-rural inequality 2010	(2) Inequality over time, 2005–2010	(3) Rural inequality over time, 2005–2010
Household size (number of members)	✓	✓	✓
Family composition (% of children under 14)	✓	✓	✓
Gender (male household head)	✓	✓	✓
Age of household head (years)	✓	✓	✓
Marital status of household head (in a couple or not)	✓	✓	✓
Education (highest level of household head/spouse)	✓	✓	✓
Climate shock (at least one climate shock)	✓	✓	✓
Health shock (at least one health shock)	✓	✓	✓
Location (urban or rural)	Not included because defines decomposition groups	✓	Not included because sample restricted to rural households
Isolation (time to food market is one hour or more)	✓	N/A*	N/A*
Security level (self-perception score, 1–5 scale)	✓	✓	✓
Electricity (% of households with electricity in community)	Not included because closely proxies for urban location	✓	✓
Transportation (at least one means of transport)	Not included because "time to food market" is included	✓	✓
Crop area (land area under cultivation)	Not included because proxies for rural location	Not included because "urban/ rural" is included	✓
Regional effects (location by province)	✓	✓	✓

*It was not possible to include the variable "time to food market" in the decompositions over time because of the large number of missing values in EPM 2005.

Determinants of Urban-Rural Inequality in 2010

In 2010, urban households had a significantly higher average consumption expenditure than rural households, and this was true across all quintiles. As illustrated in table 2.4, on average, the consumption of the poorest urban households was 125.5 percent higher than the consumption of the poorest rural households, due to both differences in endowments and differences in returns. The divergence between rural and urban consumption levels is greater for the higher quintiles and reaches 157 percent for the top quintile. Also shown are the counterfactual differences due to disparities in endowments versus disparities in returns, holding the other constant: If the bottom quintile of households in rural areas had had the same endowments as those in urban areas, urban consumption in the bottom quintile would have been only 19.8 percent higher than rural consumption in the bottom quintile, due purely to the effect of changes on the returns to such endowments, all else being equal, instead of 125.5 percent.[3] Similarly, if the bottom quintile of households in rural areas had had the same returns to their endowments as those in urban areas, urban consumption would have been 88.1 percent higher than rural consumption. Taken together, the combination of the better endowments and better returns

TABLE 2.4: Differences in Consumption Expenditure between Urban and Rural Households, and Broad Decomposition into Endowment versus Returns (2010)

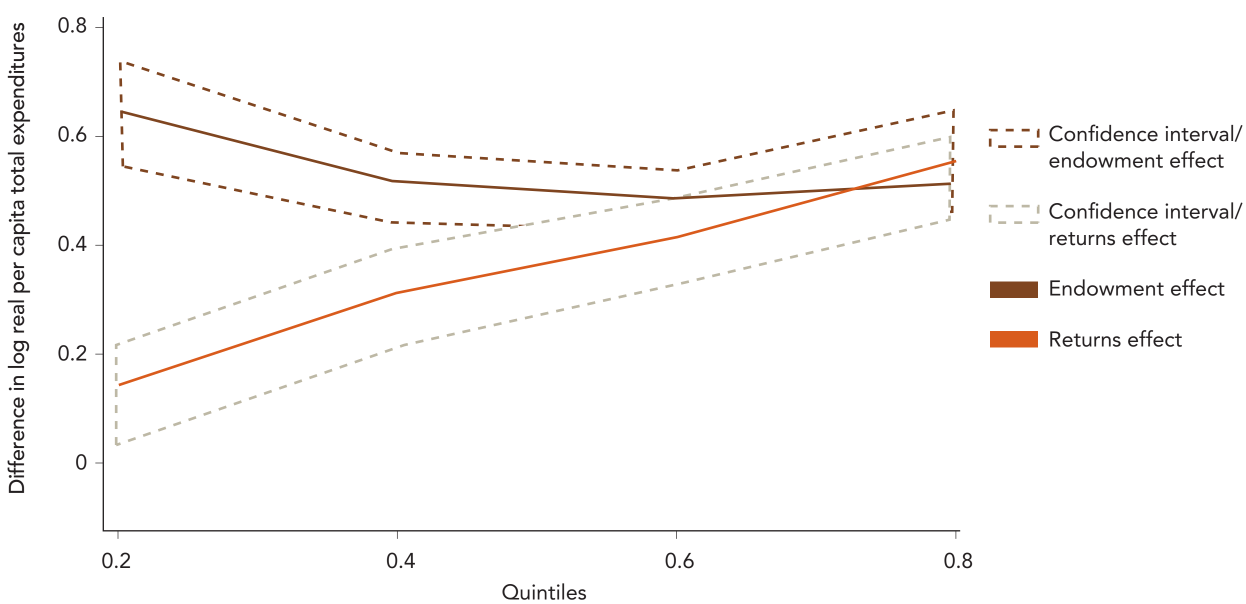
Percentiles		20	40	60	80
Overall	(Log) Consumption Urban	12.112***	12.527***	12.867***	13.257***
	(Log) Consumption Rural	11.298***	11.646***	11.936***	12.311***
	Consumption Difference	125.5%***	141.3%***	153.5%***	157.5%***
	Endowment Component	88.1%***	71.4%***	64.5%***	58.7%***
	Returns Component	19.8%***	40.8%***	54.0%***	62.3%***

Source: Calculated using EPM 2010.
Note: Percentages indicate counterfactual differences (holding other factors constant). ***Significance at the 1 percent level. **Significance at the 5 percent level. *Significance at the 10% level. Robust standard errors are used. Full results can be found in annex 2C.

enjoyed by urban households renders urban consumption in the bottom quintile over twice as high as rural consumption in the bottom quintile. Translated into shares of the underlying disparities, figure 2.6 shows the share explained by the endowment versus returns effects as estimated. Differences in endowments explain 78 percent of total consumption difference between the bottom quintile of urban and rural populations 2010, while differences in returns explain the remaining 22 percent. Toward the upper end of the distribution, the role of returns becomes more prominent. Among the urban and rural households in the top quintile, about half of the consumption gap (49 percent) is explained by differences in endowments, and the other half by differences in returns (51 percent).

Comparing households in the bottom quintiles across milieu (rural versus urban), almost three quarters of the consumption difference is attributable to differences in household size,[4] household composition, human capital, climate shocks, and distances to food markets, as shown in table 2.5. The impacts of these differences diminish for the higher quintiles, however. Across the board, urban households had higher per capita consumption due in part to their size and composition; however, by construction the welfare indicator used—per capita

FIGURE 2.6: Endowment and Return Effects between Urban and Rural Households (Percentage of Total Explained Divergence)



Source: EPM 2010.

TABLE 2.5: Contribution to Urban-Rural Consumption Differences, percentage of Total Difference, Selected Endowments (Largest Significant Contributions for Bottom Quintile)

Quintiles	20	40	60	80
Household size	13.4%	10.9%	6.2%	5.5%
Percentage of children under 14	7.4%	7.5%	10.5%	11.2%
Education level of household head	31.5%	26.9%	25.4%	23.3%
At least one climate shock	8.7%	6.5%	2.2%	1.0%
Time to food market is one hour or more	12.9%	5.7%	4.8%	4.4%
Total contribution to consumption difference	73.8%	57.4%	49.1%	45.3%

Source: Calculations using EPM 2010.

consumption—overstates differences in actual welfare, as it does not adjust for adult equivalence (within household shared resources and differential needs by age.) Apart from these demographic variables, differences were explained by urban households' having better-educated household heads. The educational attainment of the household head accounted for between 23 percent and 32 percent of total urban-rural inequality for all quintiles. Urban households were also subject to fewer climate shocks and enjoyed shorter distances to food markets than rural households did in 2010, as illustrated in table 2.6. For the bottom quintile, apart from household size the next most important correlate with consumption disparities was the time it takes to reach an urban center. While only 18 percent of urban households are located one hour or more away from the closest food market, about 62 percent of rural households are. This difference explained 12.9 percent of urban-rural consumption inequality among the poorest quintile, but significantly lower shares for richer quintiles. All together, these characteristics contributed to explaining almost three-fourths of the total difference in consumption between the urban and rural bottom quintiles.

Climate shocks also played a key role in explaining the urban-rural consumption gap. After remoteness, climate shocks were the next most important factor explaining differences in consumption between the poorest rural households and the poorest urban households. While 55 percent of the bottom quintile of rural households experienced at least one climate shock in 2010, only 23 percent of their urban counterparts did, and this difference explained about 8.7 percent of the total consumption difference between the two groups in 2010 (table 2.5).

The analysis also allows us to estimate counterfactual differences in consumption levels—that is, the degree to which rural consumption would approach urban consumption, if a given endowment or returns to that

TABLE 2.6: Summary Statistics for Selected Determinants of Urban-Rural Inequality (Largest Contributors to Inequality)

	Household size (Average number of members)		Age structure (Average members under 14)		Education level (Avg highest level completed by head/ spouse)*		Climate shocks (Households that had 1+ climate shocks)		Distance to market (Households 1h + away from food market)	
Quintile	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Bottom	5.6	6.2	46%	54%	1.9	1.4	23%	55%	18%	62%
Second	4.7	5.6	37%	49%	2.4	1.6	13%	41%	9%	53%
Middle	4.0	4.9	33%	44%	2.5	1.7	7%	37%	8%	51%
Fourth	3.6	4.3	25%	38%	2.9	1.8	7%	35%	7%	47%
Top	2.9	3.4	14%	24%	3.1	2.2	5%	27%	4%	33%

Source: EPM 2010.
*1= No schooling; 2 = primary schooling; 3 = secondary schooling; 4 = higher level.

endowment were equal across the groups. The results are presented in tables 2.7 and 2.8. If returns and all other factors were held the same for urban and rural populations, better endowments in education for the urban population in the bottom quintile would raise their consumption only 29.2 percent higher than that of the rural bottom quintile (table 2.7), instead of the actual 125 percent. However, urban households also have better returns to education, and doubles the contribution of education to the urban-rural consumption gap: if urban and rural households all had the same education levels, higher returns to education in urban areas would still render urban consumption in the bottom quintile 28.7 percent higher than the rural level (table 2.8).

Closer access to markets and reduced exposure to climate shocks made the bottom urban quintile significantly better off than the rural poorest in 2010. A greater proximity to food markets (assuming the same returns for all households) also raised urban consumption in the bottom quintile 11.1 percent above that of the bottom rural quintile. Similarly, the more limited incidence of climate shocks among urban households rendered consumption for the poorest urban people 7.4 percent higher than for their rural counterparts—a small but significant share of the total disparity.

These findings suggest that among the attributes of Madagascar's rural and urban economies, educational attainment and remoteness are key structural correlates with long-run consumption levels and urban-rural inequality, and that climatic shocks are a major short determinant. Lower returns to education in rural areas may induce lower investment in schooling, and the lower economic viability of connecting remote areas to markets could induce some "program placement" bias in our estimations. However, since such investments were made in the relatively distant past and most likely not with perfect foresight on returns in 2010, we conclude that more of these investments in rural areas would have a positive effect on consumption and reduce urban-rural inequality, but that to fully appreciate the returns to these investments, more migration, employment, and integration with urban areas would be needed.

TABLE 2.7: Counterfactual Percentage Differences in Consumption Expenditure between Urban and Rural Households, Endowments

Percentiles		20	40	60	80
Endowment	Household size	11.5%***	10.1%***	6.0%***	5.3%***
Component	Percentage of children under 14	6.2%***	6.8%***	10.3%***	11.2%***
	Education level of head/spouse	29.2%***	26.7%***	26.6%***	24.6%***
	At least one climate shock	7.4%***	5.9%***	2.0%	0.9%
	Time to food market is one hour or more	11.1%***	5.1%***	4.6%**	4.3%*
	Security level	1.8%***	1.2%***	0.4%	–0.1%

Source: Calculated using EPM 2005, 2010.
Note: Percentages indicate counterfactual differences. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in annex 2C.

TABLE 2.8: Counterfactual Percentage Differences in Consumption Expenditure between Urban and Rural Households, Returns

Percentiles		20	40	60	80
Returns	Household size	–24.0%***	–16.9%***	0.4%	–2.6%
Component	Education level of head/spouse	28.7%***	18.3%***	11.5%***	–5.9%
	At least one climate shock	–6.8%**	–6.9%***	–3.1%	–0.8%
	Security level	–14.4%**	–11.8%**	–2.3%	4.6%

Source: Calculated using EPM 2005, EPM 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in the annex 2C.

Determinants of Changes in Consumption and Inequality between 2005 and 2010, National Sample

From 2005 to 2010, the real consumption of the Malagasy households in the bottom quintile fell by 3.1 percent. Conversely, the consumption of households in the top quintile grew by 10.1 percent. Thus, inequality increased over the period, with the consumption gap between the bottom quintile and the top quintile growing by 17 percent. The results of the decomposition of changes in consumption for all households between 2005 and 2010 are presented for each quintile in tables 2.9 and 2.10. Across the distribution, endowments themselves improved. Of those significantly related to consumption levels in the RIF regressions, the endowments that helped boost consumption (or offset consumption losses) for some segments of the distribution in 2010 relative to 2005 were (1) an increase in education of the household head, (2) expanded access to electricity (in the top three quintiles),[5] and (3) greater ownership of transportation assets (table 2.11).

The net decrease in consumption for households in the bottom quintile was caused by a large drop in returns to their endowments (assets and circumstances) and a more severe experience of shocks (figures 2.7 and 2.8). Even as endowments increased for all quintiles over the period, shifts in "returns" to these factors reduced consumption for the bottom quintile and increased it for the top quintile. For the other two quintiles, there was no significant change. If households in the bottom quintile had had the same endowments in 2005 and 2010, their consumption would have fallen by 6.9 percent over the period due purely to the decline in returns. While some moderate improvements in endowment levels were observed, these were not sufficient to offset the deterioration in returns. In fact, consumption would have increased by 4.0 percent in the bottom quintile if returns had remained constant over the period. The net effect of a slight improvement in endowments and a considerable deterioration of returns was the observed net decline in consumption for the poorest (3.1 percent) (table 2.9). The primary negative shift was in returns (table 2.10).

A large portion (33.3 percent) of the (modest) improvement in household endowments was due to a reduced frequency of climate shocks.[6] While in 2005 over 59 percent of households in the bottom quintile had been affected by at least one climate shock, this proportion had dropped slightly to 52 percent by 2010. If the severity of climate shocks had been the same in 2005 and 2010, consumption in the bottom quintile would

TABLE 2.9: Counterfactual Percentage Changes in Consumption Expenditure between 2005 and 2010

Percentiles		20	40	60	80
Overall	(Log) 2010 consumption	11.390***	11.764***	12.100***	12.587***
	(Log) 2005 consumption	11.422***	11.752***	12.072***	12.491***
	(Log) cons. difference	–3.1%**	1.1%	2.9%**	10.1%***
	Endowment component	4.0%***	2.8%***	3.5%***	4.8%***
	Returns component	–6.9%***	–1.7%	–0.5%	5.0%***

Source: Calculated using EPM 2005, 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Full results can be found in Annex D.

TABLE 2.10: Endowment and Returns Components as a Percentage of Total Change in Consumption (by Quintile, All Households)

Percentiles	20	40	60	80
Endowment component	122%	255%	117%	49%
Returns component	–222%	–155%	–17%	51%

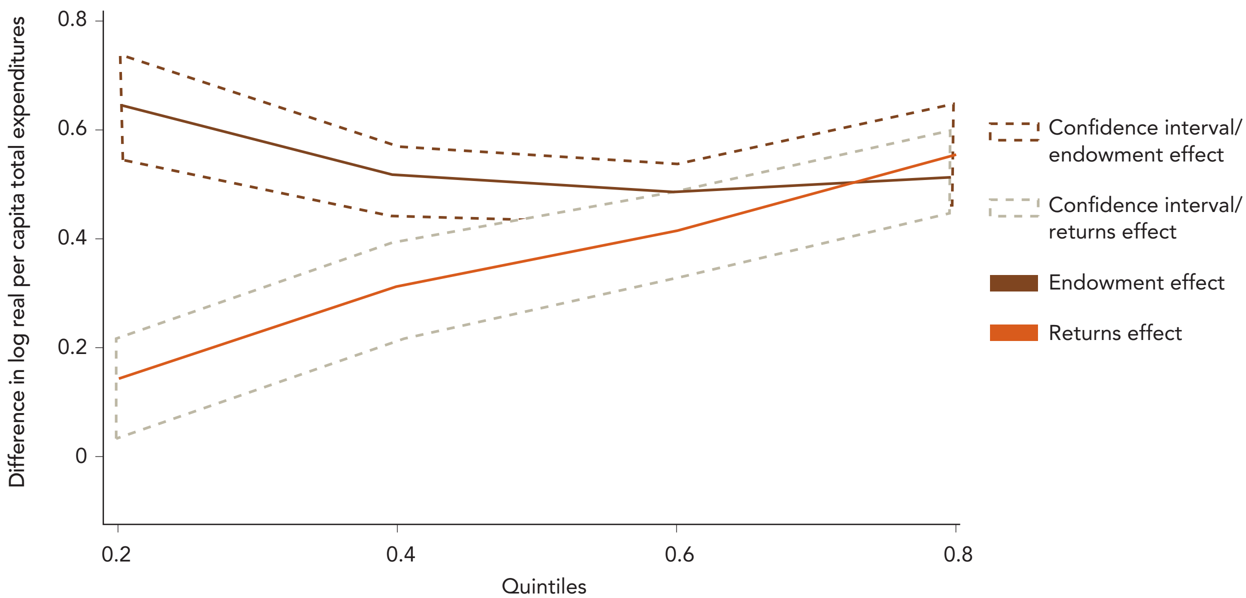
Source: Calculated using EPM 2005, 2010.

TABLE 2.11: Counterfactual Percentage Changes in Consumption Expenditure between 2005 and 2010, Changes in Endowments (Holding Returns Constant)

Percentiles	20	40	60	80
Located in rural area	0.2%**	0.2%**	0.2%**	0.3%**
Age of household head	–0.1%**	–0.1%**	–0.1%*	0.0%
Education level of head/spouse	0.6%***	0.7%***	0.9%***	1.4%***
At least one climate shock	1.3%***	0.4%	0.4%	–0.2%
Security level	0.7%***	0.5%***	0.5%**	0.6%**
Access to electricity in community	0.3%***	0.7%***	1.5%***	2.6%***
Means of transport	0.5%***	0.5%***	0.6%***	0.6%***

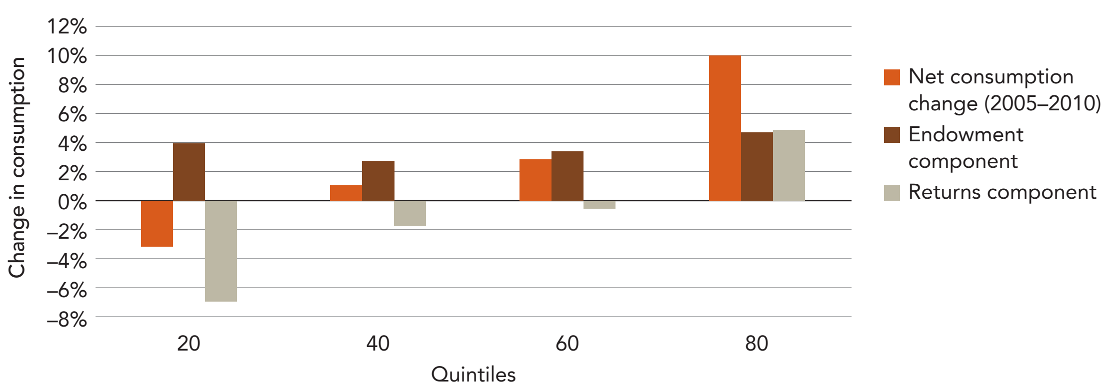
Source: Calculated using EPM, 2010.
Note: percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in annex 2D.

FIGURE 2.7: Endowment and Return Effects from 2005 to 2010 (Share of Explained Divergence, by Consumption Quintile)



Source: EPM 2005, 2010.

FIGURE 2.8: Changes in Consumption, Endowments, and Returns Components (2005–2010, by Quintile)



Source: EPM 2005, 2010.

have actually increased by 1.3 percent holding all other factors constant (table 2.11).

Despite their slightly lower frequency, climate shocks had a stronger negative impact on consumption for the poorest quintile of households in 2010 than they did in 2005, as shown in table 2.12. If they had been hit by the same *number* of shocks as in 2005, the consumption of the poorest households would have declined by as much as 5.3 percent due only to the greater severity of the 2010 shocks. Thus, compared to 2005, significantly more devastating climate shocks affected just a slightly lower proportion of households in 2010, and this explained a large share of the observed net drop in consumption among the poorest households.[7]

In addition, a large portion of the drop in consumption expenditure was explained by decreasing returns to economic activities in rural areas. Holding endowments constant, lower returns to assets and circumstances facing the rural population would account for a 5.7 percent drop in consumption for the poorest quintile between 2005 and 2010. More severe effects of health shocks also explain a significant, but smaller share of the decline in consumption (table 2.12). However, returns do not change over the period for basic household characteristics—in particular, household size, percentage of household members that are children, the age of the household head, marital status of the household head, and his or her educational attainment level. Moreover, we do not find evidence of a change in the returns to security, households' ownership of a means of transportation, or community-level access to electricity.

Over this period, however, the returns to opportunities for male-headed households relative to female-headed ones diverged for all quintiles. Holding endowments constant over the period, the poorest quintile would have experienced an increase in consumption of about 13.8 percent to due the differential benefits of having a male household head, all else equal. They would be higher for the top, by 14.2 percent, and somewhat less disparate in the middle (ranging from 7.4 to 8.5 percent) (table 2.12). As discussed below, this appears to indicate that female-headed households were less able to offset declining returns through secondary off-farm employment, and when employed, they received a lower wage.

For the bottom quintile of the distribution, the combined decrease in consumption caused by the deterioration in returns to rural activities and to more severe climate and health shocks were larger than the improvement brought about by a moderate improvement in endowments and by higher returns to male-headed households. Together these influences yielded a net 3.1 percent decrease in consumption expenditure between 2005 and 2010 for the bottom quintile. For other quintiles, returns to rural location decreased even more than for the top, but the effects of climate shocks and health shocks were less severe (table 2.12). Coupled with significant improvements in endowments (particularly education and access to electricity), this resulted in consumption levels that were either not statistically different or that were higher than the 2005 ones for the other quintiles in our analysis.

Tables 2.13 and 2.14 show the percentage of the total changes in consumption over the period associated with each significant shift in "endowments" versus the "effects" or returns to those endowments. As shown, in order of importance for shifts in consumption of the bottom 40 percent of the distribution are the change in incomes for male-headed households (with a positive effect, noted by "+"), followed by the change in incomes

TABLE 2.12: Counterfactual Percentage Changes in Consumption Expenditure between 2005 and 2010, Returns (Holding Endowments Constant)

Percentiles	20	40	60	80
Located in rural area	–5.7%**	–9.2%***	–15.5%***	–21.1%***
Male household head	13.8%***	8.5%**	7.4%*	14.2%**
At least one climate shock	–5.3%***	–2.3%*	–0.5%	1.1%
At least one health shock	–1.7%***	–1.2%***	0.0%	–1.0%

Source: Calculated using EPM 2005, EPM 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in the annex 2D.

TABLE 2.13: Contribution and Offsetting Factors to 2005–2010 Change in Consumption by Quintile, Percentage of Total Change, Selected Returns (Largest Significant Contributions for Bottom Quintile)

Percentiles	20	40	60	80
Change in returns to being located in rural area	–184%	–882%	–583%	–247%
Change in effects of climate shocks	–169%	–209%	–17%	11%
Change in effects of health shocks	–53%	–109%	0%	–10%
Change in returns to male household head	403%	745%	245%	139%

Source: Calculated using EPM 2010, EPM 2005.

TABLE 2.14: Contributions and Offsetting Factors 2005–2010 Change in Consumption, Percentage of (Absolute Value of) Total Change, Selected Endowments (Significant Effects Only)

Percentiles	20	40	60	80
Located in rural area	6%	18%	7%	3%
Age of household head	–3%	–9%	–3%	0%
Education level of household head/spouse	19%	64%	31%	15%
At least one climate shock	41%	36%	14%	–2%
Security level	22%	45%	17%	6%
Percentage of households with electricity in community	9%	64%	52%	27%
Ownership of means of transport	16%	45%	21%	6%

Source: Calculated using EPM 2010, 2005.

in rural areas (–), the effect of climate shocks (–), effects of health shocks (–), the education of the household head (+), the reduced frequency of climate shocks (+), the level of community electrification (+), the security level (+), and ownership of some means of transport (+). Figure 2.9 shows the effects of the factors that were significant for the bottom quintiles—climate and health shocks, returns to rural economic activities, and differential gains by male-headed households.

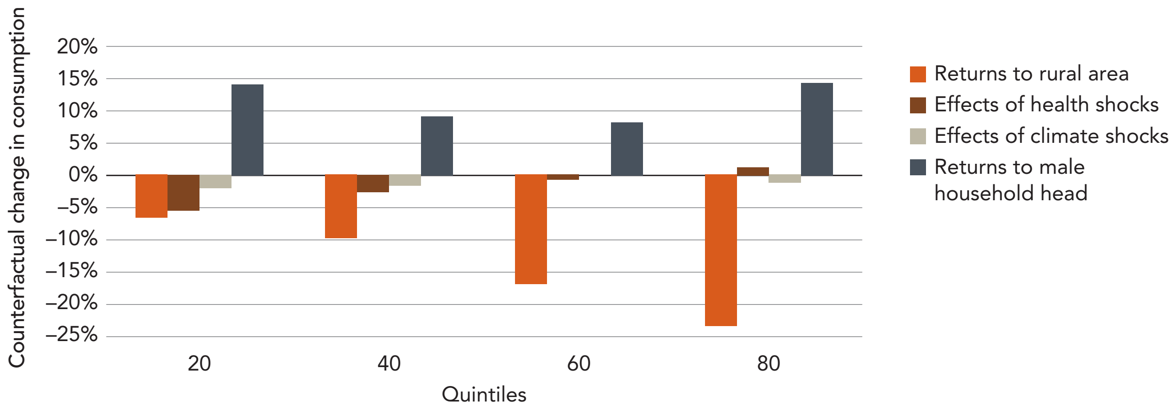
Determinants of Changes in Consumption and Inequality between 2005 and 2010, Rural Households Only

To better understand the reasons for declines in returns to rural economic activity and analyze factors that may affect rural households differently, we next decompose changes over time for rural households only.

From 2005 to 2010, the consumption of rural households in the bottom quintile of the distribution fell by 6.0 percent, almost twice as much as the drop experienced by the poorest rural and urban households together, 3.1 percent (table 2.15). In contrast, the consumption of the top rural quintile did not change significantly. Results of the decomposition of changes in consumption between 2005 and 2010 among rural households are presented for each quintile in tables 2.16 and 2.17.

As expected given the results of the foregoing section "Determinants of Change in Consumption and Inequality between 2005 and 2010, National Sample," the drop in consumption for the bottom rural quintile was explained for the most part by a fall in returns (figure 2.10 and figure 2.11). Modest improvements in endowments were not sufficient to offset the decrease in consumption brought about by these falling returns. Holding returns constant, rural households would have had a 2.8 percent higher in consumption in 2010 than in 2005. However, holding endowments constant, returns would have caused consumption to drop by 8.6 percent over the same period (table 2.15).

FIGURE 2.9: Main Determinants of Change in Consumption (2005–2010)



Source: EPM 2005, 2010.
Note: Effects smaller than 2% or not significant for the bottom quintile not pictured.

The increased severity of climate shocks in 2010 was the main determinant of the decline in consumption for the bottom rural consumption quintile. Madagascar is highly susceptible to cyclones, floods, droughts, locust infestations, and animal and plant diseases, which expose the population to considerable risks. Lacking adequate mechanisms for ex ante or ex post risk mitigation, the population of Madagascar is particularly vulnerable to climate risks, which can cause a great deal of physical destruction and erode the livelihoods of the rural population, in particular (Auffret 2014).

Between 2005 and 2010, Madagascar was hit by a series of particularly severe climate shocks, which caused extensive physical damage and led to widespread food insecurity. In 2008, three consecutive cyclones hit the country, affecting 17 out of 22 regions (Auffret 2014). In 2010, the southern regions were affected by prolonged droughts, which had a devastating impact on harvests. During the same year, the cyclone Hubert brought considerable damage to eastern coastal provinces and generated severe floods, which destroyed large quantities of agricultural production. As a result, food insecurity issues affected over 80 percent of the Malagasy population in 2010 (FAO 2010).

Whereas the percentage of rural households in the bottom quintile experiencing at least one climatic shock declined slightly, from over 60 percent in 2005 to 55 percent 2010, the adverse effects of climatic shocks in 2010 were greater than in 2005. The effect of the fall in the frequency of experiencing these shocks would have caused a small increase in consumption for the poorest rural quintile, had the severity (or "returns") held constant with those in 2005 (+1.8 percent) (table 2.16). However, consumption would have

TABLE 2.15: Influences on Changes in Consumption Expenditure between 2010 and 2005, Endowments versus Returns Component (Percentage of 2005 Consumption, Rural Only)

Percentiles		20	40	60	80
Overall	(Log) rural 2010 consumption	11.303***	11.652***	11.945***	12.323***
	(Log) rural 2005 consumption	11.365***	11.677***	11.964***	12.319***
	Consumption change	–6.0%***	–2.5%**	–1.9%	0.4%
	Endowment component	2.8%***	1.3%*	0.5%	1.3%
	Returns component	–8.6%***	–3.7%***	–2.3%*	–0.9%

Source: Calculated using EPM 2005 and EPM 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Full results can be found in Annex E.

TABLE 2.16: Contributing and Offsetting Influences on Net Changes in Consumption between 2005 and 2010, Rural Only: Endowments (Percentage Changes, Significant Effects Only)

Percentiles		20	40	60	80
Endowments	Household size	–0.6%*	–0.6%*	–0.6%*	–0.5%*
	Percentage of children under 14	–0.3%**	–0.4%***	–0.6%	–0.9%
	Education level of head/spouse	0.4%***	0.5%***	0.5%	0.7%
	At least one climate shock	1.8%***	0.5%	–0.1%	0.0%
	Security level	0.6%***	0.2%	0.2%	0.4%
	Access to electricity in community	0.4%***	1.1%***	1.6%	2.9%
	Means of transportation	0.3%***	0.4%***	0.4%	0.5%
	Cultivated land	–0.5%***	–0.5%***	–0.5%	–0.6%

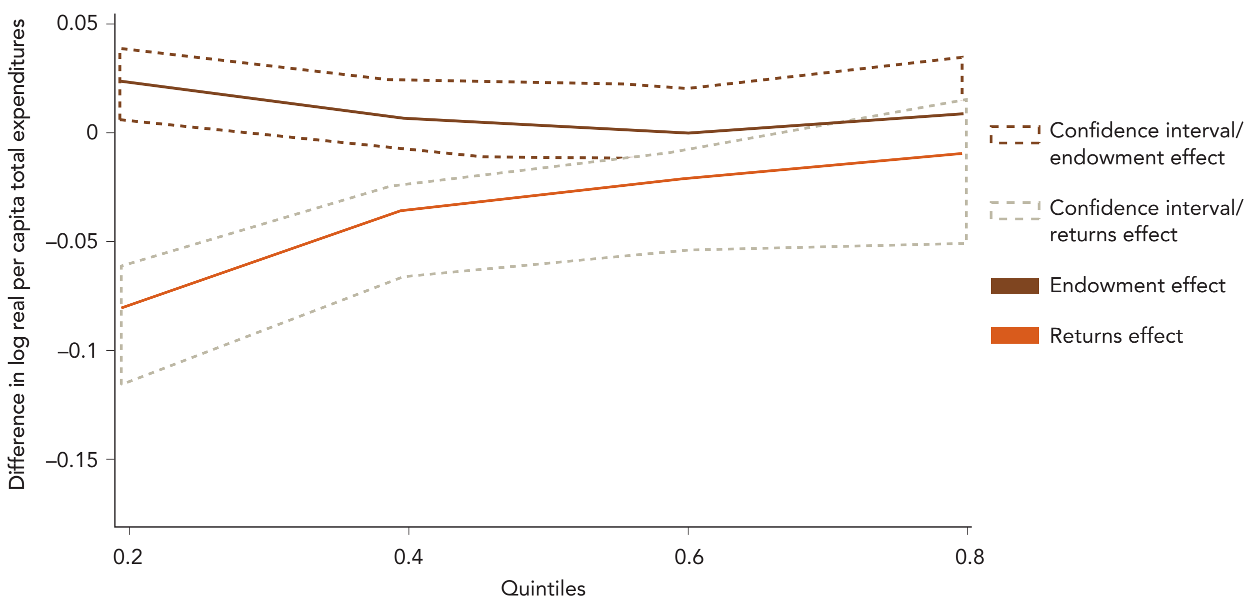
Source: Calculated using EPM 2005, 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in annex 2C.

TABLE 2.17: Contributing and Offsetting Influences on Net Changes in Consumption between 2005 and 2010, Rural Only, Returns (Percentage Changes, Significant Effects Only)

Percentiles		20	40	60	80
Returns	Male household head	18.1%***	15.4%***	9.5%**	7.1%
	At least one climate shock	–7.0%***	–2.3%	0.9%	0.3%
	At least one health shock	–2.0%***	–1.4%**	0.2%	0.9%
	Access to electricity in community	–0.4%*	0.1%	–0.2%	0.0%
	Cultivated land	–6.4%***	–6.1%***	–6.3%***	–3.8%**

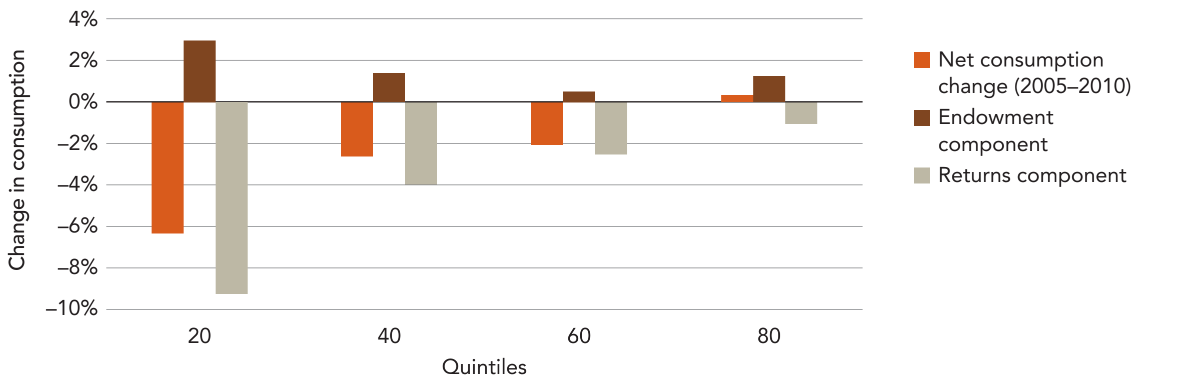
Source: Calculated using EPM 2005, 2010.
Note: Percentages indicate counterfactual changes. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Robust standard errors are used. Only determinants significant for the bottom quintile are presented. Full results can be found in annex 2E.

FIGURE 2.10: Endowment and Return Effects from 2005 to 2010 (Rural Households)



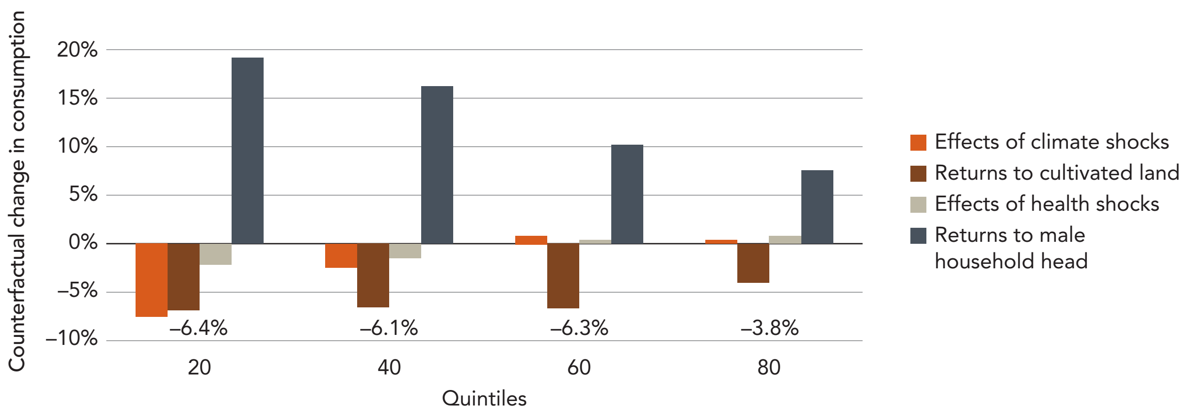
Source: EPM 2005, 2010.

FIGURE 2.11: Changes in Consumption, Rural Households, Endowments and Returns Components (2005–2010)



Source: EPM 2005, 2010.

FIGURE 2.12: Main Determinants of Change in Consumption (Rural Households, 2005–2010)



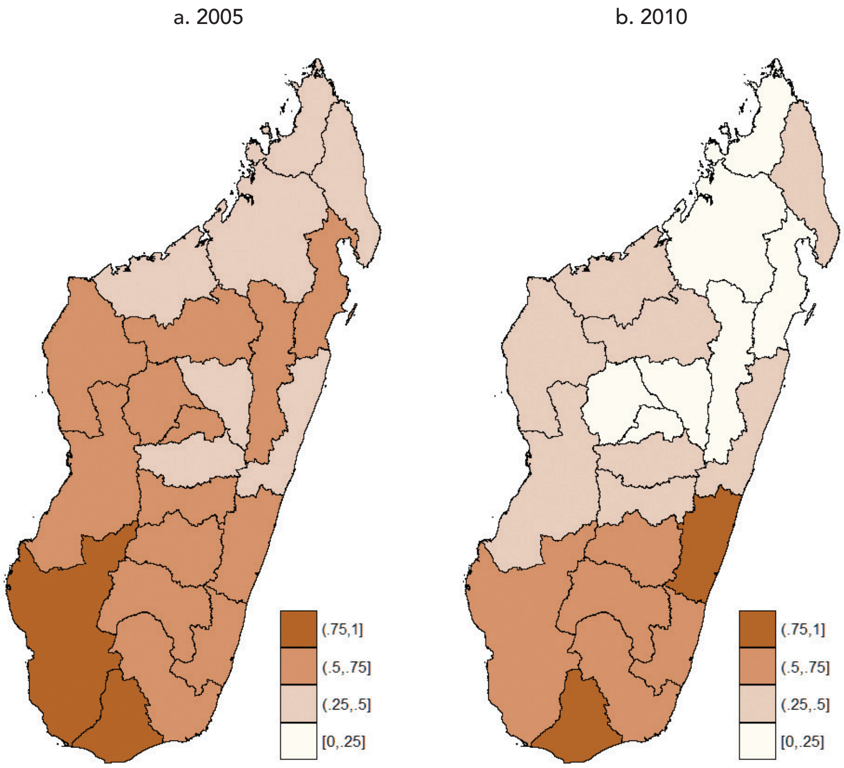
Source: EPM 2005, 2010.
Note: Effects smaller than 2% or not significant for the bottom quintile not pictured.

decreased by 7.0 percent due to the stronger effects of climate shocks alone (holding endowments constant) in 2010 (table 2.17, figure 2.11). The negative effects of climate shocks on the bottom rural quintile were significantly more pronounced than those observed for all households combined (–5.3 percent change, as reported for the national sample in the foregoing section). Rural households depend for their livelihoods on subsistence agriculture and are more vulnerable to climatic events, and therefore such shocks explain a greater portion of the change in consumption over time for rural households than for the national sample.

Although different households experience shocks in different years, there is some spatial correlation in the frequency of reported climatic shocks.

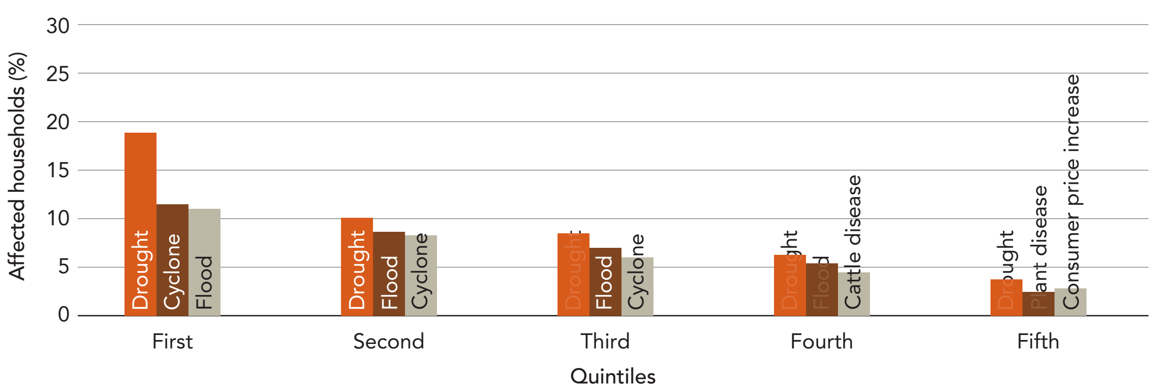
Figure 2.13 shows the proportion of households affected by at least one climate shock by province in 2005 and 2010, and figure 2.14 shows the top three shocks for 2010. Although not perfectly correlated, one of the southern regions with a high percentage of households affected in 2005 is also highly affected in 2010, and shocks are more frequent in the south and the west in both years. Droughts, cyclones, floods, late rains, plant diseases, locust invasions, and cattle diseases were frequently reported shocks in all consumption quintiles, but they affected a significantly greater share of the poorest households than households in other quintiles. Droughts, cyclones, and floods were the three most-cited shocks in the bottom three quintiles (more so than any other type of economic, health, or security shock). Almost one in five households in the bottom quintile (18.6 percent)

FIGURE 2.13: Proportion of Households Hit by at Least One Climate Shock (by Province)



Source: Calculated using EPM 2005, 2010.
Note: Numbers in brackets represent the proportion of households, with 1 = 100 percent.

FIGURE 2.14: Top Three Shocks by Quintile (2010)



Source: EPM 2010.

reported being hit by drought in 2010, while only 10.3 percent in the second quintile did so, 8.5 percent in the third quintile, 6.3 percent in the fourth quintile, and 3.8 percent in the top quintile. The same pattern was true for cyclones: 11.5 percent of the poorest households were hit by cyclones, but only 8.2 percent in the second quintile, 6.0 percent in the third quintile, 3.7 percent in the fourth quintile, and 2.2 percent in the top quintile. Finally, floods were experienced by 10.9 percent of households in the bottom quintile, 8.7 percent in the second quintile, 7.2 percent in the third quintile, 5.4 percent in the fourth quintile, and only 2.0 percent in the

top quintile. This, combined with the RIF decomposition results, shows the key role that climatic shocks played in deepening poverty in 2010. As both the frequency and severity of these shocks are expected to increase due to climate change, the vulnerability of the poor population to extreme weather events is also likely to increase. Over the next 50 to 100 years, average temperatures in Madagascar are expected to rise by 2.5 degrees. As a consequence, average annual rainfall in Madagascar is forecast to decrease, while at the same time sharp increases in precipitation will occur during the rainy season (Auffret 2014).

Changes in household and community assets also played a role. Households in the bottom two quintiles had greater means of transportation, greater community-level access to electricity, and greater education relative to the bottom two quintiles in 2005. Offsetting this was a decline in the area of land cultivated (table 2.16).

As with the national sample, a decline in returns to these factors explains shifts between the two years more than asset accumulation. A significant decline in returns to cultivated land explains a large part of the decrease in consumption for rural households in the bottom quintile of the distribution between 2005 and 2010 (table 2.17). If endowments had not varied over the period, the poorest rural households would have experienced a drop of 6.4 percent in their consumption due to a decline in returns to the land they cultivate. Middle quintiles experienced drops in returns to cultivated land of a similar magnitude, but for the top quintile the drop was considerably smaller.

As was the case for the national sample, increased returns to male-headed households relative to female-headed households prevented average consumption in the bottom quintile from dropping even further. Holding endowments constant, male-headed households would have experienced an 18.1 percent increase in consumption between 2005 and 2010, which exceeds the effect in the national sample (rural and urban households combined of 13.8 percent) just discussed, rather than 6 percent.

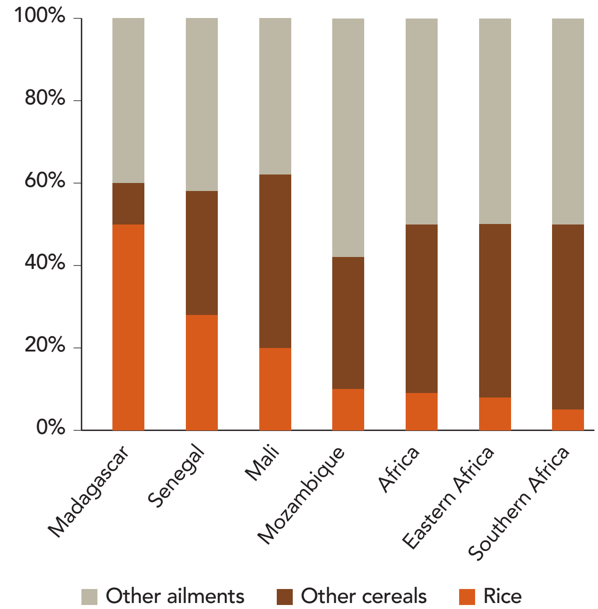
Explaining Changing Patterns in Returns

In this section, we provide additional analysis of the underlying reasons for falling returns to land and to living in rural areas between 2005 and 2010, as well as of the differential returns to male-headed households over time. The key factors identified are policies in rice markets and deteriorating transport conditions, which worsened the already problematic degree of integration of Madagascar's rice markets (Moser, Barrett, and Minten 2009). In addition, outcomes in labor markets suggest that preferences for males in the off-farm labor market in a year of low agricultural productivity made a more significant difference in households' ability to offset agricultural losses through off-farm work.

RICE MARKETS AND POLICIES

Most Malagasy households are not only highly dependent on agriculture for their livelihoods but are especially dependent on rice. Rice constitutes a more important proportion of their consumption than it does for households in other countries in Sub-Saharan Africa (figure 2.15), and the majority of them are both consumers and producers of the grain. Rice paddy is produced by 87 percent of agricultural producers and is the main crop across all quintiles of the income distribution.

FIGURE 2.15: Food Supply by Type of Aliment (kcal/capita/day)



Source: FAOSTAT 2016.

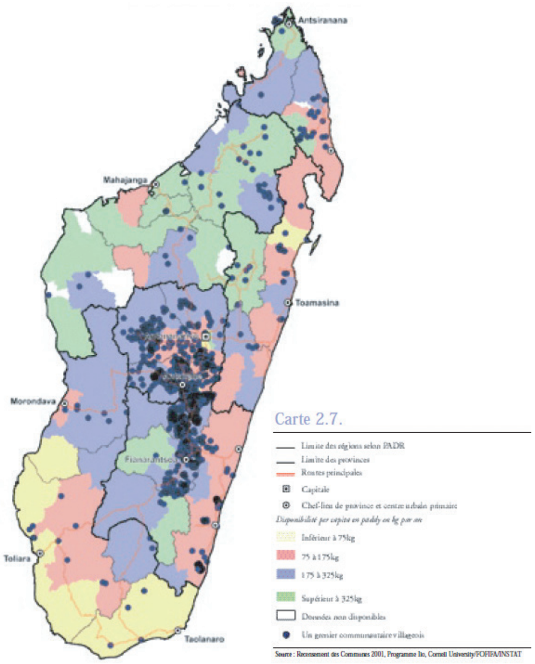
Richer households tend to be slightly more concentrated in rice production, with almost 60 percent of their total production being rice paddy relative to 40 percent for the bottom quintile (table 2.18. Given

TABLE 2.18: Share of Each Product in Total Production by Consumption Quintile (2005)

	Quintile of consumption				
Group of products	1 (poorest)	2	3	4	5 (richest)
Paddy	39.9	46.0	49.3	54.6	59.9
Maize and other	2.2	2.4	2.7	2.2	2.1
Cassava	28.4	26.7	22.4	21.2	17.3
Sweet potatoes	6.4	7.1	6.3	5.5	4.1
Other tubers	2.6	4.1	5.4	4.0	4.0
Leguminous	2.1	2.2	2.3	2.2	2.0
Vegetable	3.0	1.9	1.9	2.2	1.8
Fruit	7.8	5.2	5.2	4.2	4.5
Industrial culture	6.1	3.5	3.6	3.2	3.3
Cash crops	1.4	1.0	0.9	0.8	0.9
Total	100.0	100.0	100.0	100.0	100.0

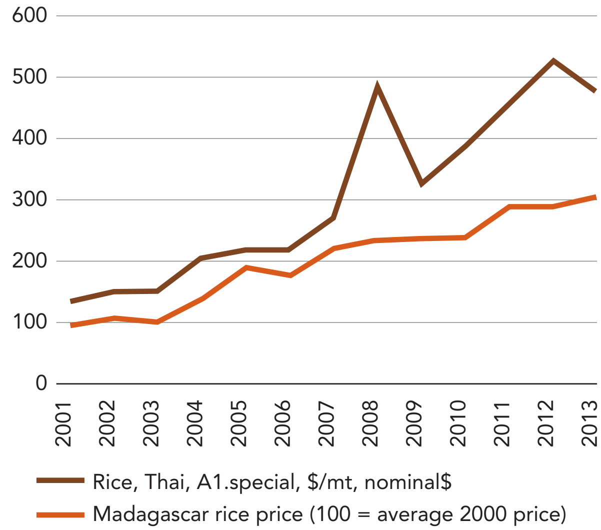
Source: EPM 2005.

FIGURE 2.16: Location of Community Granaries



Source: Randrianarisoa 2003.

FIGURE 2.17: Movements in World and Local Rice Prices



Source: World Bank Commodity Price databank and INSTAT.

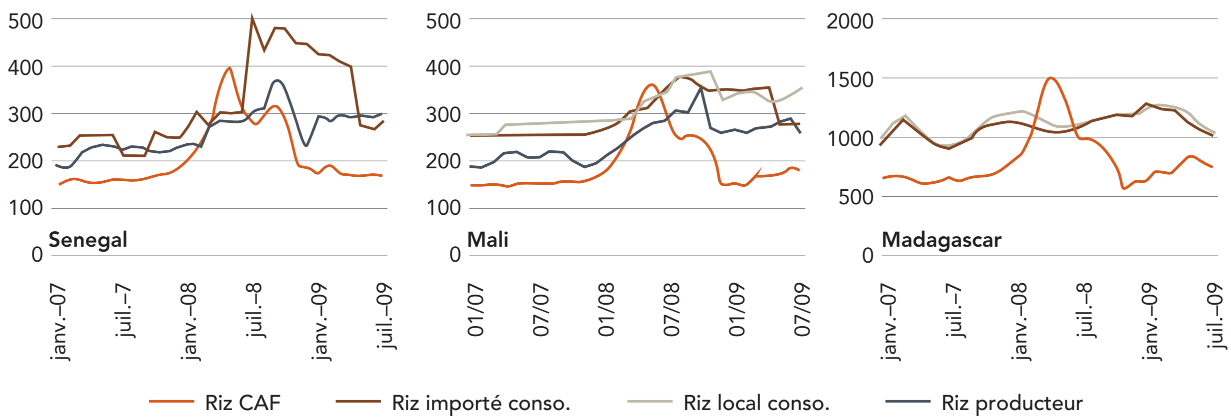
the scarcity of adequate storage facilities (figure 2.16), most Malagasy households are not able to economically store sufficient rice for self-consumption over extended periods of time. They tend to alternate between being net sellers and net buyers of rice in different seasons of the year. Large shares of the rural population become net rice consumers: Among the 68 percent of households that produce rice, more than two-thirds need to purchase rice at some point during the year (Auffret 2014). Many also experience severe seasonal food shortfalls. In 2005, higher rice prices coincided with relatively good rice yields to improve the incomes of rice producers, particularly net rice sellers. In fact, poorer farmers opted to sell more of their rice crop than richer farmers, who may have preferred to consume more of it (see table 2.19). The international price of rice rose significantly after 2005,

TABLE 2.19: Share of Crop Sold by Consumption Quintile (2005)

	Quintile of consumption				
Group of products	1 (poorest)	2	3	4	5 (richest)
Paddy	30.8	28.4	27.8	28.1	23.7
Maize and other	37.3	37.6	30.6	38.0	36.6
Cassava	29.2	31.7	30.2	28.2	32.6
Sweet potatoes	27.6	22.0	20.2	21.8	28.6
Other tubers	35.7	36.3	35.7	42.8	36.9
Leguminous	47.5	49.6	48.6	52.3	50.4
Vegetable	73.3	54.4	58.1	53.1	56.1
Fruit	57.6	58.2	58.7	57.1	52.7
Industrial culture	25.7	41.8	46.9	48.3	44.9
Cash crops	81.2	87.5	83.5	75.5	79.5

Source: EPM 2005.

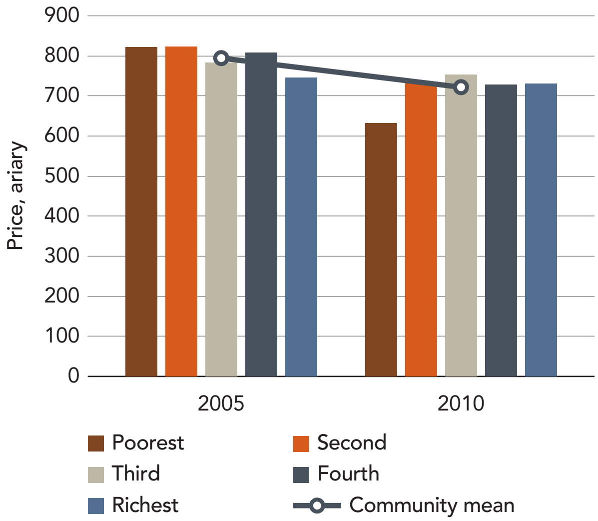
FIGURE 2.18: International and Domestic Rice Prices, Selected Countries (2007–2009)



Source: David-Benz and Lancon 2013.

spiking in 2008 and 2012; however, this did not fully translate into increased revenues for Malagasy rice producers. The government of Madagascar sought successfully to prevent a more rapid escalation of domestic rice prices in the face of rising world prices, and the domestic price of rice was kept relatively stable (figure 2.17). In 2007, the government removed tariffs on rice imports and decreased *ad valorem* taxes on rice, removing value-added tax on rice imports completely in July of 2008. Anticipating drought and further increases in the international rice price the government preordered rice imports (50,000 metric tons of Indian rice) and banned rice exports (David-Benz 2011). As a result, Madagascar's consumer rice price increases were milder than those of several other Sub-Saharan rice importer countries (figure 2.18).

FIGURE 2.19: Median Nominal Price Received for Rice Paddy (by Consumption Quintile)



Source: EPM 2005, 2010.

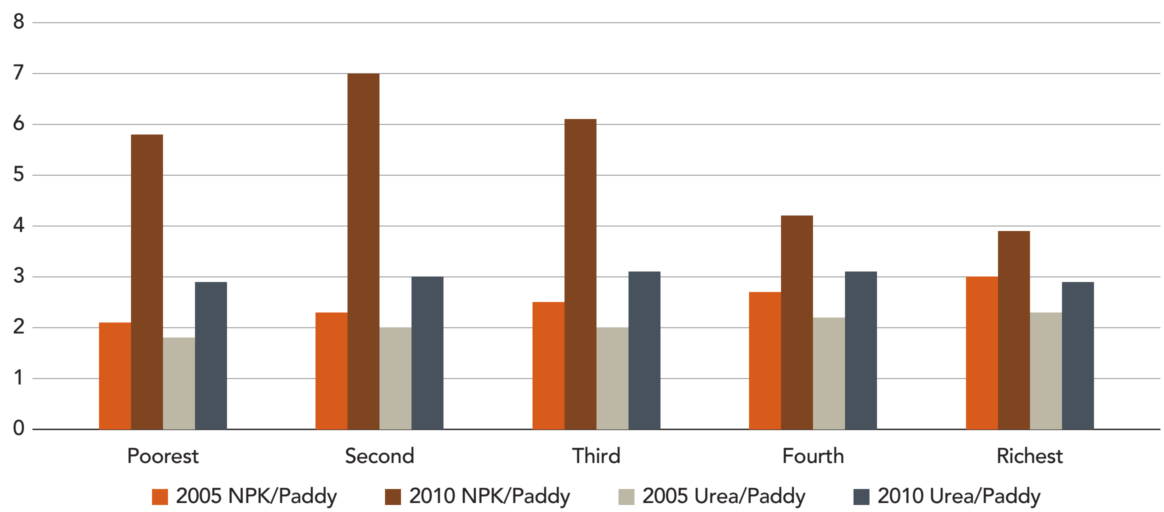
In fact, prices paid to rice producers were in many cases significantly lower in 2010 than they had been in 2005; the poor in particular received a much lower price. Figure 2.19 shows that the median producer price of rice was higher in 2005 at 546 ariary per kilogram than in 2010, at 533 ariary, despite continued increases in both the world price and the domestic market price in urban centers.[8] This nominal decline coincided with a cumulative increase in the consumer price index of 58 percent.

Responding to the worsening terms of trade, across all quintiles, there was a shift out of rice relative to 2005 (table 2.20), and poorer households sold a smaller share of their crop in 2010, whereas richer households with more advantageous market access increased this share (table 2.21).

Although lower prices improve the welfare of net consumers, including those in urban areas, they are detrimental to producers. In a variety of Asian countries, for example, high food prices have been found to push up rural wages and reduce poverty rates—in the long term and in some cases in the short term, despite the adverse impact on some net food buyers (Ivanic and Martin 2014). Indeed, in a period of rising rice prices in Madagascar, between 2001 and 2005, real consumption among the bottom 40 percent grew (see Belghith, Osborne, and Randriankolona 2016).

On net, relative to 2005, producers in 2010 experienced a decline in the terms of trade in agriculture in rural areas, especially those in the bottom three quintiles (figure 2.20). While local producer prices generally fell, the price of agricultural inputs increased considerably, and the correlation between adverse terms of trade in agriculture and poverty is pronounced in 2010 in a way that it was not in 2005. These factors are likely to partly explain the decrease in returns to cultivated land and to being located in a rural area for the bottom quintile.

FIGURE 2.20: Relative Prices of Fertilizer versus Rice (2005 and 2010 by Household Consumption Quintile)



Source: EPM 2005, 2010. NPK = Nitrogen, Phosphorus, Potassium.

TABLE 2.20: Share of Each Production in Total Production by Consumption Quintile (2010)

	Quintile of consumption				
Group of products	1	2	3	4	5
Paddy	35.8	40.0	44.9	45.0	47.2
Maize and others	3.8	3.8	3.7	3.8	5.4
Cassava	33.6	24.6	23.1	18.5	16.7
Sweet potatoes	8.0	6.7	4.5	5.9	3.8
Other tubers	1.4	2.2	2.0	2.7	1.9
Leguminous	2.2	2.3	2.9	2.6	2.8
Vegetable	1.9	3.1	2.9	4.5	4.2
Fruit	7.0	10.4	8.5	10.1	9.1
Industrial culture	4.5	5.5	6.3	5.6	7.5
Cash crops	1.8	1.3	1.2	1.1	1.3
Total	100.0	100.0	100.0	100.0	100.0

Source: EPM 2010.

TABLE 2.21: Share of Crop Sold by Consumption Quintile (2010)

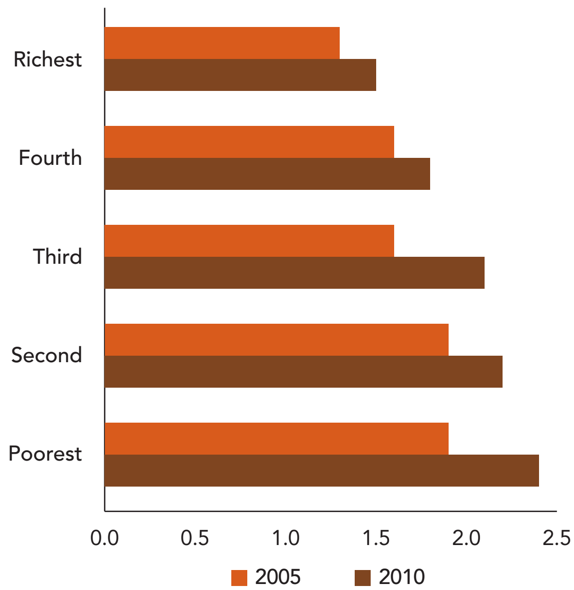
	Quintile of consumption				
Group of products	1	2	3	4	5
Paddy	24.8	24.0	26.5	27.2	33.2
Maize and others	38.0	42.4	42.5	47.8	62.3
Cassava	31.5	36.6	42.2	39.6	47.8
Sweet potatoes	25.2	23.0	18.5	24.8	32.3
Other tubers	17.1	26.6	33.9	38.1	37.0
Leguminous	53.8	57.9	61.0	59.1	65.3
Vegetable	60.4	75.8	69.5	75.3	84.1
Fruit	60.6	76.5	67.8	64.3	73.4
Industrial culture	31.9	39.8	37.3	40.4	62.8
Cash crops	82.2	86.8	83.8	83.6	84.5

Source: EPM 2010.

THE ROLE OF INFRASTRUCTURE

In addition to rice policies, sharp rises in transport costs due to deteriorating physical infrastructure appear to have affected market integration and the returns and endowments of rural households.[9] Between 2005 and 2010, while the terms of trade worsened in agriculture, particularly for the poor, the poor also became significantly more isolated (figure 2.21). The time required to reach various markets and services—already high in 2005—increased dramatically between 2005 and 2010. A sharp increase in transport costs, affecting all quintiles, is observed between 2005 and 2010, with the highest transport costs observed for households in the bottom quintile. This signals a deterioration of road connectivity as a consequence of the deep political crisis of 2009 (after which revenues for road maintenance diminished). For example, for the poorest quintile in 2005, the closest primary school was 44 kilometers away on average, whereas this doubled to 90 kilometers by 2010. The time needed to reach a food market rose for the bottom quintile from 1.9 to 2.4 hours, and the time to reach a main urban center from less than 6 to almost 12 hours. Also striking from the figures is that in 2010 consumption is closely and inversely related to these distances and times to reach schools, food markets, and urban centers, much more so than in 2005. For example, whereas it took over 90 minutes for the richest quintile to reach food markets,

FIGURE 2.21: Average Time to Food Market by Quintile (Hours)



Source: EPM 2005, 2010.

it took over 144 minutes for the poorest quintile.[10] In addition, the cost of transporting goods to the nearest main urban center increased dramatically between 2005 and 2010. As shown in table 2.22, the average costs by region increased by between 36 and 80 percent in the rainy season over this timeframe, and up to 99 percent for Antsirinana in the dry season. Because it is unlikely that the deterioration in connectivity could have been closely correlated with poverty rates, particularly in a time of postcrisis scarcity in fiscal and other resources, it is likely that this deterioration had a causal effect on poverty. By increasing the cost of accessing inputs and reducing the producer price for agricultural households, while also making access to health services more difficult, deteriorating transport conditions reduced the real incomes and consumption levels of those most affected.

Access to electricity also appears to play a role in stimulating a modest increase in local incomes. For the average rural household in the poorest quintile, in 2005 less than 1 percent of other households in the community were connected to the electric grid. By 2010, this proportion had not changed significantly. In contrast, rates of access to electricity increased considerably for other groups: the proportion of community households with electricity went from about 12 percent to about 18 percent for the top rural quintile, and from 40 percent to 48 percent for the top urban quintile. As illustrated in figure 2.22, little progress had been made by 2010 in reaching areas outside of the province of the capital city with electricity, and this may have arguably contributed to deepening the divide between urban and rural areas and between top and bottom quintiles. At the same time, over the period 2005 to 2010, it was in communities with the poorest households (in the bottom two quintiles) where the rate of community-level electricity access was associated with higher consumption. This pattern does not fit the standard story: that income determines electrification. While the decline in returns to electricity for the bottom quintile completely offsets the gain in the electrification rate, for the second quintile, this is not the case. On the margin, greater access to electricity appears to help increase economic opportunities, albeit only modestly.

RETURNS BY GENDER

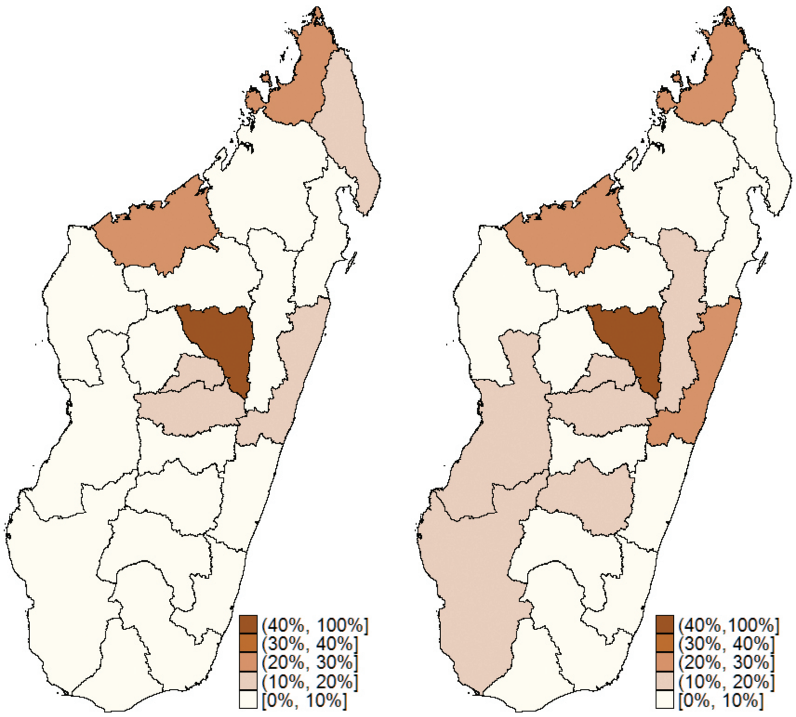
All in all, between 2005 and 2010, the poorest Malagasy households were severely affected by a combination of severe climate shocks, lower producer rice prices, increasing agricultural input costs, rising transport costs and deteriorating connections to markets. This caused their consumption to decline significantly, often leading to food insecurity, particularly in rural areas. In this analysis, this is evidenced by strong decreases in returns to being located in a rural area, to cultivated land, and by large negative effects of climate shocks.

TABLE 2.22: Transport Cost for a 50-Kilogram Bag of Rice, Dry Season and Rainy Season Average (in 2005 U.S. Dollars, by Quintile)

Quintile	Bottom	Second	Middle	Fourth	Top	All
2005	$1.65	$1.57	$1.41	$1.32	$1.00	$1.40
2010	$2.19	$2.08	$2.07	$1.96	$1.66	$1.99

Source: EPM 2005, 2010.

FIGURE 2.22: Average percentage of Households with Electricity in a Community (Percentage Ranges by Province)



Source: EPM 2005, 2010.

However, the analysis also shows sharp increases in returns to male-headed households relative to female-headed ones, with the divergence greater for poorer quintiles. While these increases have not been sufficiently high to completely offset the negative impacts of falling returns to cultivated land and agriculture, nor the devastating effects of climate shocks, they have prevented the poorest households headed by males from falling much deeper into poverty, particularly in rural areas.

A possible explanation is that men were better able to find employment outside of the agricultural sector between 2005 and 2010, while women were unemployed or stayed in agriculture, experiencing falling returns. As shown in figure 2.23, between 2005 and 2010, the proportion of individuals looking for employment increased sharply but was considerably higher among females. Men were significantly better able to find nonfarm employment and/or to find a second job than women over the period, as shown. Moreover, the gap in wages between males and females appears higher in 2010 than for other years, for at least some prime working ages (approximately age 40), as shown in figure 2.24.

Another possible contributing explanation is that the returns to activities in which males are more likely to engage increased relative to those for females. For the average 40-year-old worker, the wage disparity in 2010 between females and males was higher than in other years (figure 2.24). Bi and Osborne (2016) also find that among urban-based owner-occupied microenterprises, male entrepreneurs have higher returns than female ones, even controlling for other factors. Sectors such as transport or logging tend to be disproportionately pursued by males, and these sectors may have experienced increased profitability relative to those females pursued disproportionately, such as textiles production.

Narrowing the focus to the group that suffered the largest consumption decline over the period of interest (i.e., households in the poorest rural quintile), we find further support for this hypothesis. Agriculture remained the main sector of employment for both male and female household heads in the poorest rural quintile both in 2005 and 2010. As shown in figure 2.25: Employment of Household Heads in the Bottom Quintile in Rural Areas (by Sector), the vast majority of male household heads (92.4 percent) and female household heads (79.6 percent) were employed in agriculture in 2005. By 2010, this proportion had decreased only slightly: to 90.8 percent for males and to 76.4 percent for females, respectively. Despite falling returns to agriculture and severe climate shocks, it appears that the overwhelming majority of household heads in the poorest rural quintile did not shift their main source of livelihood away from the primary sector, arguably due to a lack of alternative opportunities.

An examination of employment patterns for household heads in the bottom quintile of the distribution of rural households by gender can shed light on the increase in

FIGURE 2.23: Labor Market Outcomes, 2005, 2010, and 2012 (Kernel-Weighted Local Polynomial Smoothed Age-Wage Profiles

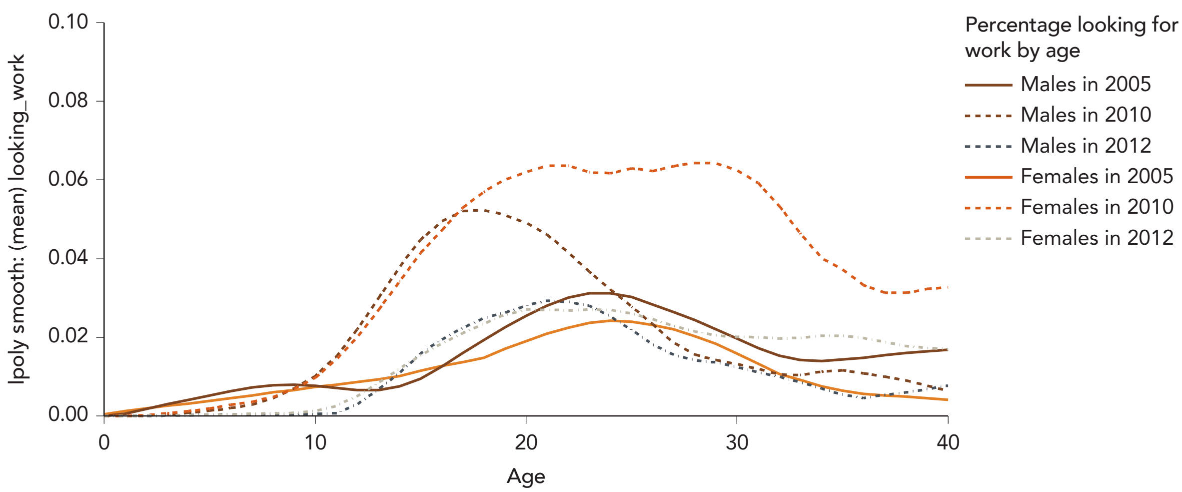
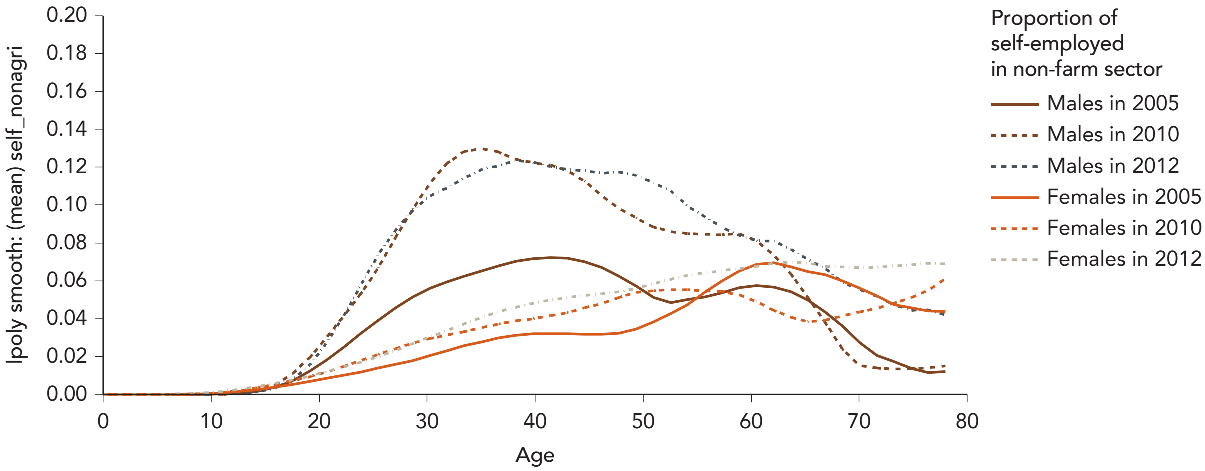


FIGURE 2.23: Labor Market Outcomes, 2005, 2010, and 2012 (Kernel-Weighted Local Polynomial Smoothed Age-Wage Profiles *(continued)*



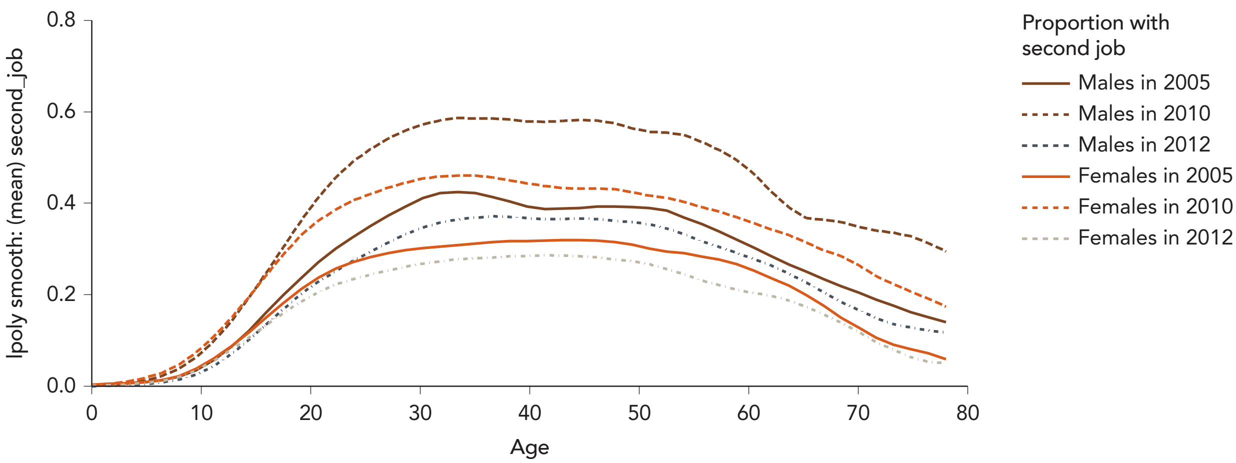
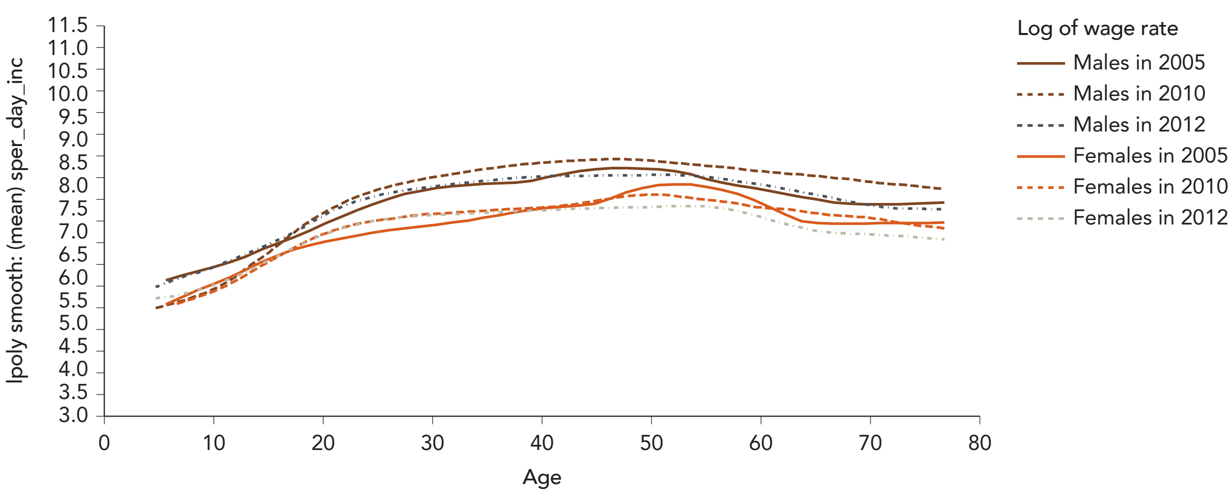
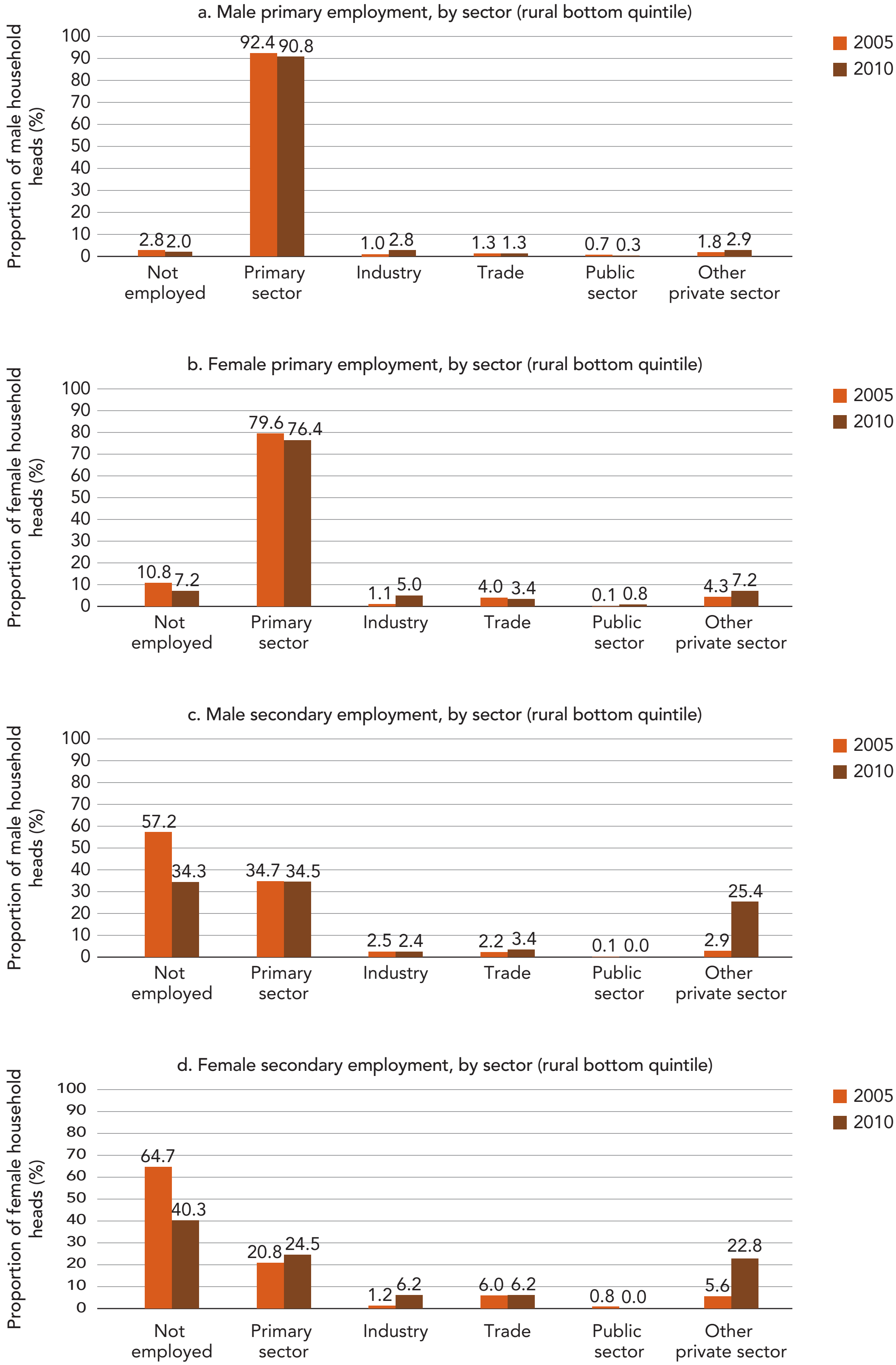


FIGURE 2.24: Log of Wage, Male and Female Workers by Age (2005, 2010, and 2012)



Sources EPM 2005, 2010, and ENSOMD 2012.
Notes: x axis = age of worker. y axis represents the smoothed polynomial of log wage per day.

FIGURE 2.25: Employment of Household Heads in the Bottom Quintile in Rural Areas (by Sector)



Source: EPM 2005, 2010.

gender-based differences in consumption over the period. A large proportion of these household heads (both male and female) took on secondary activities between 2005 and 2010. Whereas only 42.8 percent of male household heads in the bottom quintile in rural areas had secondary employment in 2005, by 2010 this proportion had reached 65.7 percent (figure 2.25). Among their female counterparts, a similar trend was observed: in 2005, whereas only 35.3 percent of female household heads had secondary employment, this rose to 59.7 percent in 2010. Thus, the increases were similar—at 22.9 percent for males and 24.4 percent for females. Many poor rural household heads took on secondary activities alongside their primary employment in order to cope with falling returns to agriculture (their main source of livelihood) and the devastating effects of climate shocks.

The main gender-based differences revealed are that for these males almost all new secondary employment was in the "other private sector," which absorbed 22.5 percent of them. For female household heads in this population segment, however, only 17.2 percent were able to find secondary employment in the "other private" sector, with the other 7.2 percent finding secondary work in the primary sector and industry. It is likely that males engaged in more remunerative activities than their female counterparts, such as transport and other nontrade services, and this made them better able to cope with falling returns to agriculture and climate shocks. Because between 2005 and 2010 most poor rural household heads remained primarily employed in agriculture, a large share of both male and female household heads sought and found secondary employment. Because secondary activities in the private sector accounted for most of the new employment for both genders, there are likely to be additional explanations for the sharp increase in returns to male household heads observed over the period. The increase in returns to male household heads in the bottom quintile could be due to the fact that the opportunities available mostly to males were more remunerative than those available mostly to females, on average.

Annex 2A.

MAP 2A.1: Administrative Divisions of Madagascar

a. Old division

b. New divisions as of 2009

Source: Wikipedia.

Annex 2B. Methodology

The first part of this analysis investigated sources of inequality between rural and urban households. The second part explored the factors influencing the evolution of inequality in consumption over time. In this section, the methodology behind these decompositions is presented in greater detail.

Inequality is measured using real monthly household per capita consumption expenditures adjusted for spatial and temporal variations in the cost of living. The consumption aggregate includes expenditures on both food and nonfood items and excludes both rental housing and durable goods expenses. The aggregate is constructed following the methodology suggested by Deaton and Zaidi (2002). The adjustment for price variations across regions and over time is done using the Fisher index of unit values from the surveys.

To analyze the sources of inequality between rural and urban groups and over time, the unconditional quantile regression method is applied. This method allows to understand how the differences in the distributions of observed household characteristics between groups or over time contribute to the welfare gap. It also identifies how the marginal "effects" of these characteristics vary across the entire distribution. Overall, the method allows to distinguish between the contributions of: (a) differences in household characteristics ("endowment" effects); and, (b) disparities in returns to these characteristics ("returns" effect) to inequality.

The development of decomposition methods has been a fertile area of research over the last few decades. Building on the seminal work of Oaxaca (1973) and Blinder (1973), several procedures that allow one to go beyond the mean have been put forward and have been used widely in the literature. Popular approaches used in the decomposition of distributional statistics and the analysis of the sources of inequality include the standard Oaxaca–Blinder decomposition method, the reweighting procedure of DiNardo, Fortin, and Lemieux (1996) and the quantile-based decomposition approach of Machado and Mata (2005). The main drawback of the Oaxaca–Blinder technique is that it applies the decomposition only to the mean welfare differences between two population subgroups and yields an incomplete representation of the inequality sources. The other conventional methods extend the decomposition beyond the mean and permit the analysis of the entire distribution. Nevertheless, they all share the same shortcoming in that they involve several assumptions and computational difficulties (Fortin, Lemieux, and Firpo 2010).

The RIF-regression method proposed by Firpo, Fortin, and Lemieux (2009) addresses these shortcomings and allows one to evaluate the impact of changes in the distribution of the explanatory variables on quantiles of the unconditional (marginal) distribution of the outcome variable. To distinguish this approach from other conditional quantile regressions (Koenker and Bassett 1978; Koenker 2005), this method is referred to as *unconditional quantile regression*. It allows one to decompose the welfare gaps at various quantiles of the unconditional distribution into differences in households' endowment characteristics such as education, age, employment status, and so forth, and differences in the marginal (conditional) correlations between consumption and these characteristics. These components are then further decomposed to identify the specific attributes which contribute to the widening welfare gap.

The procedure is carried out in two stages. The first stage consists of estimating unconditional quantile regressions on log real per capita consumption for group 1 and group 2 households, then constructing a counterfactual distribution that would prevail if group 1 households had received the returns that pertained to group 2. The comparison of the counterfactual and empirical distributions allows us to estimate the part of the welfare gap attributable to households' characteristic differentials, the endowment effect, and the part explained by differences in returns to those characteristics, the return effect. The second stage involves dividing the endowment and return components into the contribution of each specific characteristic variable.

The method can be easily implemented as a standard linear regression, and an ordinary least squares (OLS) regression of the following form can be estimated:

$$RIF(y, Q_\theta) = X\beta + \varepsilon, \tag{1}$$

where y is log real per capita monthly household consumption and $RIF(y,Q_\theta)$ is the RIF of the θth quantile of y estimated by computing the sample quantile Q_θ and estimating the density of y at that point by the kernel method:

$$RIF(y,Q_\theta) = Q_\theta + \left(\theta - I\left\{y \le Q_\theta\right\}\right) / f_y(Q_\theta),$$

where f_y is the marginal density function of y and I is an indicator function. RIF can be estimated by replacing Q_θ by θth sample quantile and estimating f_y by kernel density. X is the regressors matrix including the intercept, β is the regression coefficient vector, and ε is the error term.

We estimate the model for each decile from the 10th to 90th quantiles and use the unconditional quantile regression estimates to decompose the rural-urban inequality, as well as the changes in consumption between 2005 and 2010 into a component attributable to differences in the distribution of characteristics and a component due to differences in the distribution of returns. This is done as follows:

$$\hat{Q}_\theta^i - \hat{Q}_\theta^{i'} = \left\{\hat{Q}_\theta^i - \hat{Q}_\theta^*\right\} + \left\{\hat{Q}_\theta^* - \hat{Q}_\theta^{i'}\right\} = \left(\bar{X}^i - \bar{X}^{i'}\right)\hat{\beta}_\theta^i + \bar{X}^{i'}\left(\hat{\beta}_\theta^i - \hat{\beta}_\theta^{i'}\right), \tag{2}$$

where $\hat{Q}_\theta$ is the θth unconditional quantile of log real per capita monthly household consumption, $\bar{X}$ represents the vector of covariate averages and $\hat{\beta}_\theta$ the estimate of the unconditional quantile partial effect. Superscripts i, i', and * designate, respectively, the urban (or 2010), rural (or 2005), and counterfactual values. The first term on the right-hand side of equation (2) represents the contribution of the differences in distributions of household characteristics to inequality at the θth unconditional quantile, denoted endowment effect. The second term of the right-hand side of the equation represents the inequality due to differences (or discrimination) in returns to the household characteristics at the θth unconditional quantile. The endowment and return effects can be further decomposed into the contribution of individual specific household characteristics (or group of some characteristics) as follows:

$$\hat{Q}_\theta^i - \hat{Q}_\theta^* = \sum_k \left(\bar{X}_k^i - \bar{X}_k^{i'}\right)\hat{\beta}_{\theta,k}^{\ i} \quad \text{and} \quad \hat{Q}_\theta^* - \hat{Q}_\theta^{i'} = \sum_k \bar{X}_k^{i'}\left(\hat{\beta}_{\theta,k}^{\ i} - \hat{\beta}_{\theta,k}^{\ i'}\right) \ k{:}1\ldots K, \tag{3}$$

where k designates the individual specific household characteristics.

$\hat{Q}_\theta^* = X^{i'}\hat{\beta}^i$ is the counterfactual quantile of the unconditional counterfactual distribution which represents the distribution of welfare that would have prevailed for group i' (rural/2005 households) if they have received group i (urban/2010 households) returns to their characteristics. The decomposition results may vary with the choice of the counterfactual distribution. For example, if the counterfactual used is the distribution that would have prevailed for group i if they have received group i' returns we would obtain different results. The choice of the counterfactual in this analysis is motivated by the aim of emphasizing household groups living in disadvantaged areas.

Annex 2C. Determinants of Urban-Rural Inequality in 2010

TABLE 2C.1: Summary Statistics for Urban and Rural Households (2010)

	Household size (Average number of members)		Age structure (Average members under 14)		Gender of head (Households with male head)		Age of head (Average years)		Marital status of head (Household heads with spouse)	
Quintile	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**
Bottom	5.6	6.2	46%	54%	72%	78%	42	41	72%	76%
Second	4.7	5.6	37%	49%	78%	82%	43	41	75%	80%
Middle	4.0	4.9	33%	44%	79%	82%	41	42	72%	79%
Fourth	3.6	4.3	25%	38%	76%	84%	41	42	69%	79%
Top	2.9	3.4	14%	24%	75%	80%	43	43	59%	70%

	Education level (Avg highest level completed by head/ spouse, 1–4)		Health shocks (Households that had 1+ health shocks)		Climate shocks (Households that had 1+ climate shocks)		Distance to market (Households 1+ hours away from food market)		Security level (Avg security: 1 = very poor, 4 = very good)	
Quintile	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**
Bottom	1.9	1.4	9%	7%	23%	55%	18%	62%	3.1	3.1
Second	2.4	1.6	9%	4%	13%	41%	9%	53%	2.8	3.0
Middle	2.5	1.7	8%	4%	7%	37%	8%	51%	2.7	3.0
Fourth	2.9	1.8	6%	5%	7%	35%	7%	47%	2.7	3.0
Top	3.1	2.2	4%	6%	5%	27%	4%	33%	2.7	2.9

Source: EPM 2010.

TABLE 2C.2: Decomposition of Log Consumption Expenditure, Urban and Rural Households (2010)

	Percentiles	20	40	60	80
Overall	(Log) Urban consumption	12.112	12.527	12.867	13.257
		(0.022)***	(0.020)***	(0.019)***	(0.024)***
	(Log) Rural consumption	11.298	11.646	11.936	12.311
		(0.011)***	(0.010)***	(0.010)***	(0.012)***
	Difference	0.813	0.881	0.930	0.946
		(0.025)***	(0.022)***	(0.021)***	(0.027)***
	Endowment component	0.632	0.539	0.498	0.462
		(0.037)***	(0.027)***	(0.025)***	(0.034)***
	Returns component	0.181	0.342	0.432	0.484
		(0.047)***	(0.032)***	(0.027)***	(0.031)***
N		11,820	11,820	11,820	11,820

Source: Calculated using EPM 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2C.3: Decomposition of Log Consumption Expenditure, Urban and Rural Households (2010, Endowments)

Percentiles		20	40	60	80
Endowments component	Household size	0.109	0.096	0.058	0.052
		(0.013)***	(0.010)***	(0.008)***	(0.009)***
	Percentage of children under 14	0.060	0.066	0.098	0.106
		(0.013)***	(0.011)***	(0.010)***	(0.013)***
	Male household head	–0.005	–0.002	0.001	–0.006
		(0.004)	(0.003)	(0.003)	(0.005)
	Age of household head	0.001	0.001	0.000	0.001
		(0.002)	(0.001)	(0.000)	(0.001)
	Household head has a spouse	–0.007	–0.002	0.002	0.014
		(0.005)	(0.005)	(0.005)	(0.007)*
	Education level of head/spouse	0.256	0.237	0.236	0.220
		(0.019)***	(0.017)***	(0.016)***	(0.022)***
	At least one health shock	–0.001	–0.002	–0.003	–0.003
		(0.002)	(0.002)	(0.002)*	(0.002)**
	At least one climate shock	0.071	0.057	0.020	0.009
		(0.022)***	(0.016)***	(0.013)	(0.014)
	Time to food market is one hour or more	0.105	0.050	0.045	0.042
		(0.031)***	(0.023)**	(0.022)**	(0.025)*
	Security level	0.018	0.012	0.004	–0.001
		(0.005)***	(0.004)***	(0.004)	(0.006)
N		11,820	11,820	11,820	11,820

Source: Calculated using EPM 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2C.4: Decomposition of Log Consumption Expenditure, Urban and Rural Households (2010, Returns)

Percentiles		20	40	60	80
Returns component	Household size	–0.274	–0.185	0.004	–0.026
		(0.084)***	(0.062)***	(0.051)	(0.061)
	Percentage of children under 14	–0.081	–0.036	–0.088	0.035
		(0.053)	(0.044)	(0.041)**	(0.054)
	Male household head	–0.054	–0.089	–0.130	–0.036
		(0.065)	(0.057)	(0.058)**	(0.088)
	Age of household head	0.072	–0.007	–0.052	0.087
		(0.075)	(0.065)	(0.063)	(0.088)
	Household head has a spouse	0.028	0.019	0.023	0.004
		(0.057)	(0.052)	(0.052)	(0.081)
	Education level of head/spouse	0.252	0.168	0.109	–0.061
		(0.042)***	(0.036)***	(0.035)***	(0.052)
	At least one health shock	–0.000	–0.008	–0.014	–0.018
		(0.005)	(0.004)*	(0.004)***	(0.005)***
	At least one climate shock	–0.070	–0.072	–0.032	–0.008
		(0.032)**	(0.023)***	(0.019)	(0.022)
	Time to food market is one hour or more	–0.066	–0.002	0.013	0.042
		(0.040)	(0.030)	(0.028)	(0.033)
	Security level	–0.155	–0.126	–0.023	0.045
		(0.066)**	(0.060)**	(0.060)	(0.082)
	Constant	0.003	0.002	0.016	0.017
		(0.015)	(0.011)	(0.010)	(0.012)
N		11,820	11,820	11,820	11,820

Source: Calculated using EPM 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies. Source:

Annex 2D. Determinants of Inequality between 2005 and 2010

TABLE 2D.1: Summary Statistics for All Households (2005 and 2010)

	Location (Households in rural area)		Household size (Average number of members)		Age structure (Average members under 14)		Gender of head (Households with male head)		Age of head (Average years)	
Quintile	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**
Bottom	95%	98%	6.1	6.1	52%	53%	84%	78%	42	42
Second	92%	94%	5.3	5.4	46%	47%	80%	81%	42	41
Middle	88%	89%	4.9	4.8	42%	42%	80%	82%	43	42
Fourth	79%	76%	4.1	4.1	34%	34%	79%	82%	43	42
Top	56%	47%	3.5	3.3	25%	22%	79%	78%	43	42

	Marital status of head (Households heads with spouse)		Education level (Highest level completed by head/spouse, 1–4)		Climate shocks (Households that had 1+ climate shocks)		Electricity (Avg households with electricity in community)	
Quintile	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**
Bottom	80%	77%	1.5	1.4	59%	52%	4%	2%
Second	78%	79%	1.6	1.6	57%	39%	6%	5%
Middle	77%	79%	1.7	1.8	57%	35%	9%	9%
Fourth	72%	76%	1.9	2.0	48%	27%	18%	21%
Top	71%	67%	2.4	2.6	35%	14%	40%	48%

Source: EPM 2010.

TABLE 2D.2: Decomposition of Log Consumption Expenditure (between 2010 and 2005)

Percentiles		20	40	60	80
Overall	(Log) 2010 consumption	11.390	11.764	12.100	12.587
		(0.010)***	(0.009)***	(0.010)***	(0.014)***
	(Log) 2005 consumption	11.422	11.752	12.072	12.491
		(0.010)***	(0.009)***	(0.009)***	(0.013)***
	Difference	–0.032	0.011	0.029	0.096
		(0.014)**	(0.012)	(0.014)**	(0.019)***
	Endowment component	0.039	0.028	0.034	0.047
		(0.007)***	(0.007)***	(0.009)***	(0.012)***
	Returns component	–0.071	–0.017	–0.005	0.049
		(0.014)***	(0.012)	(0.013)	(0.018)***
N		24,157	24,157	24,157	24,157

Source: Calculated using EPM 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2D.3: Decomposition of Log Consumption Expenditure between 2010 and 2005, Endowments

Percentiles		20	40	60	80
Endowments component	Located in rural area	0.002	0.002	0.002	0.003
		(0.001)**	(0.001)**	(0.001)**	(0.001)**
	Household size	–0.000	–0.000	–0.000	–0.000
		(0.003)	(0.003)	(0.003)	(0.002)
	Percentage of children under 14	–0.001	–0.002	–0.002	–0.003
		(0.001)	(0.001)	(0.002)	(0.003)
	Male household head	–0.000	–0.000	–0.000	–0.000
		(0.000)	(0.000)	(0.001)	(0.001)
	Age of household head	–0.001	–0.001	–0.001	–0.000
		(0.001)**	(0.001)**	(0.000)*	(0.000)
	Household head has a spouse	–0.000	–0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.001)
	Education level of head/spouse	0.006	0.007	0.009	0.014
		(0.001)***	(0.002)***	(0.002)***	(0.003)***
	At least one climate shock	0.013	0.004	0.004	–0.002
		(0.004)***	(0.003)	(0.004)	(0.004)
	At least one health shock	0.002	–0.003	–0.007	–0.003
		(0.003)	(0.002)	(0.002)***	(0.004)
	Security level	0.007	0.005	0.005	0.006
		(0.002)***	(0.002)***	(0.002)**	(0.003)**
	Access to electricity in community	0.003	0.007	0.015	0.026
		(0.001)***	(0.002)***	(0.003)***	(0.005)***
	Means of transport	0.005	0.005	0.006	0.006
		(0.001)***	(0.001)***	(0.001)***	(0.001)***
N		24,157	24,157	24,157	24,157

Source: Calculated using EPM 2005, 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2D.4: Decomposition of Log Consumption Expenditure (between 2010 and 2005, Returns)

Percentiles		20	40	60	80
Returns component	Located in rural area	–0.059	–0.097	–0.169	–0.237
		(0.027)**	(0.027)***	(0.036)***	(0.069)***
	Household size	0.014	–0.020	0.016	0.008
		(0.039)	(0.030)	(0.031)	(0.040)
	Percentage of children under 14	–0.018	–0.022	–0.055	–0.097
		(0.026)	(0.022)	(0.026)**	(0.036)***
	Male household head	0.129	0.082	0.071	0.133
		(0.038)***	(0.032)**	(0.038)*	(0.059)**
	Age of household head	0.031	0.016	0.017	0.008
		(0.042)	(0.036)	(0.041)	(0.057)
	Household head has a spouse	–0.023	–0.010	–0.026	–0.152
		(0.033)	(0.029)	(0.034)	(0.053)***
	Education level of head/spouse	0.019	0.020	0.044	0.098
		(0.027)	(0.024)	(0.028)	(0.044)**
	At least one climate shock	–0.054	–0.023	–0.005	0.011
		(0.016)***	(0.014)*	(0.015)	(0.018)
	At least one health shock	–0.017	–0.012	–0.000	–0.010
		(0.006)***	(0.005)***	(0.005)	(0.008)
	Security level	–0.022	–0.018	–0.004	–0.045
		(0.039)	(0.033)	(0.036)	(0.049)
	Access to electricity in community	–0.008	–0.001	0.006	0.018
		(0.007)	(0.007)	(0.010)	(0.018)
	Means of transport	–0.002	–0.003	–0.012	–0.017
		(0.005)	(0.005)	(0.006)*	(0.010)*
	Constant	–0.029	0.057	0.102	0.322
		(0.089)	(0.078)	(0.092)	(0.143)**
N		24,157	24,157	24,157	24,157

Source: Calculated using EPM 2005, 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

Annex 2E. Determinants of Rural Inequality between 2005 and 2010

TABLE 2E.1: Summary Statistics for Rural Households (2005 and 2010)

	Household size (Avg number of members)		Age structure (Avg members under 14)		Gender of head (Households with male head)		Age of head (Average years)		Marital status of head (Households heads with spouse)	
Quintile	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**
Bottom	6.2	6.2	52%	54%	84%	84%	42	41	80%	76%
Second	5.4	5.6	47%	49%	82%	82%	42	41	78%	80%
Middle	4.9	4.9	42%	44%	82%	82%	42	42	80%	79%
Fourth	4.2	4.3	36%	38%	80%	80%	43	42	74%	79%
Top	3.5	3.4	27%	24%	79%	79%	43	43	69%	70%

	Education level (Avg highest level completed by head/spouse, 1–4)		Climate shocks (Households that had 1+ climate shocks)		Electricity (Avg households with electricity in community)		Transportation (Households with 1+ means of transport)		Land (Average acres of exploited land)	
Quintile	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**	**2005**	**2010**
Bottom	1.5	1.4	60%	55%	0.9%	1.1%	9%	7%	119	93
Second	1.6	1.6	61%	41%	1.2%	2.0%	14%	14%	127	112
Middle	1.6	1.7	60%	37%	2.0%	3.8%	16%	20%	138	121
Fourth	1.7	1.8	57%	35%	3.8%	7.0%	22%	25%	138	129
Top	2.0	2.2	51%	27%	12.2%	17.5%	33%	35%	132	149

Source: EPM 2005, 2010.

TABLE 2E.2: Decomposition of Log Consumption Expenditure, Rural Households (2005 and 2010)

Percentiles		20	40	60	80
Overall	(Log) 2010 rural consumption	11.303	11.652	11.945	12.323
		(0.011)***	(0.009)***	(0.009)***	(0.012)***
	(Log) 2005 rural consumption	11.365	11.677	11.964	12.319
		(0.010)***	(0.009)***	(0.010)***	(0.012)***
	Difference	–0.062	–0.025	–0.019	0.004
		(0.015)***	(0.013)**	(0.013)	(0.017)
	Endowment component	0.028	0.013	0.005	0.013
		(0.008)***	(0.008)*	(0.008)	(0.010)
	Returns component	–0.090	–0.038	–0.023	–0.009
		(0.015)***	(0.013)***	(0.013)*	(0.017)
N		17,755	17,755	17,755	17,755

Source: Calculated using EPM 2005, 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2E.3: Decomposition of Log Consumption Expenditure, Rural Households (2005 and 2010, Endowments)

Percentiles		20	40	60	80
Endowments component	Household size	–0.006	–0.006	–0.006	–0.005
		(0.003)*	(0.004)*	(0.003)*	(0.003)*
	Percentage of children under 14	–0.003	–0.004	–0.006	–0.009
		(0.001)**	(0.002)***	(0.002)***	(0.003)***
	Male household head	0.000	0.000	0.000	0.000
		(0.001)	(0.001)	(0.001)	(0.001)
	Age of household head	–0.001	–0.001	–0.001	–0.000
		(0.001)	(0.001)	(0.001)	(0.000)
	Household head has a spouse	0.000	–0.000	–0.000	–0.001
		(0.000)	(0.000)	(0.000)	(0.001)
	Education level of head/spouse	0.004	0.005	0.005	0.007
		(0.002)***	(0.002)***	(0.002)***	(0.003)***
	At least one climate shock	0.018	0.005	–0.001	–0.000
		(0.004)***	(0.004)	(0.004)	(0.004)
	At least one health shock	0.004	–0.001	–0.007	–0.011
		(0.003)	(0.003)	(0.003)***	(0.004)***
	Security level	0.006	0.002	0.002	0.004
		(0.002)***	(0.002)	(0.002)	(0.002)*
	Access to electricity in community	0.004	0.011	0.016	0.029
		(0.001)***	(0.002)***	(0.002)***	(0.004)***
	Means of transportation	0.003	0.004	0.004	0.005
		(0.001)***	(0.001)***	(0.001)***	(0.002)***
	Land	–0.005	–0.005	–0.005	–0.006
		(0.001)***	(0.001)***	(0.001)***	(0.002)***
N		17,755	17,755	17,755	17,755

Source: Calculated using EPM 2005, 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

TABLE 2E.4: Decomposition of Log Consumption Expenditure, Rural Households (2005 and 2010, Returns)

Percentiles		20	40	60	80
Returns component	Household size	–0.000	–0.049	0.039	0.030
		(0.043)	(0.032)	(0.031)	(0.037)
	Percentage of children under 14	–0.010	–0.022	–0.038	–0.139
		(0.030)	(0.025)	(0.027)	(0.035)***
	Male household head	0.166	0.143	0.091	0.069
		(0.043)***	(0.036)***	(0.040)**	(0.057)
	Age of household head	0.054	0.034	0.000	–0.036
		(0.045)	(0.038)	(0.041)	(0.053)
	Household head has a spouse	–0.017	–0.042	–0.004	–0.052
		(0.038)	(0.032)	(0.035)	(0.051)
	Education level of head/spouse	0.037	0.040	0.038	0.071
		(0.029)	(0.025)	(0.028)	(0.038)*
	At least one climate shock	–0.073	–0.023	0.009	0.003
		(0.018)***	(0.015)	(0.016)	(0.019)
	At least one health shock	–0.020	–0.014	0.002	0.009
		(0.006)***	(0.005)**	(0.005)	(0.008)
	Security level	–0.057	–0.029	–0.006	–0.085
		(0.042)	(0.034)	(0.036)	(0.045)*
	Access to electricity in community	–0.004	0.001	–0.002	–0.000
		(0.002)*	(0.002)	(0.003)	(0.005)
	Means of transportation	–0.001	0.007	–0.013	–0.012
		(0.006)	(0.006)	(0.006)*	(0.009)
	Land	–0.066	–0.063	–0.065	–0.039
		(0.016)***	(0.013)***	(0.014)***	(0.018)**
	Constant	–0.086	–0.077	–0.110	0.076
		(0.088)	(0.076)	(0.081)	(0.105)
N		17,755	17,755	17,755	17,755

Source: Calculated using EPM 2005, 2010.
Note: Robust standard errors in parentheses. ***Significance at the 1% level. **Significance at the 5% level. *Significance at the 10% level. Controls include province dummies.

NOTES

1. Madagascar uses per capita consumption as its welfare indicator for measuring poverty (as is consistent with the World Bank's official international poverty measure). Because of economies of scale and differential consumption needs within the household, inequalities due to differences in household size and percentage of children are likely to be overstated.
2. A full explanation of all causes of a distribution or its changes would also require an examination of broader policy, institutional, and contextual factors, which are not observable at this level of data.
3. All percentages indicate counterfactual changes in relative consumption levels between the two groups, meaning that all other factors are held equal to their values at baseline (rural group or year 2005).
4. Madagascar uses per capita consumption as its welfare indicator for measuring poverty (as is consistent with the World Bank's official international poverty measure). Because of economies of scale and differential consumption needs within the household, inequalities due to differences in household size and percentage of children are likely to be overstated.
5. Access to electricity declined in the bottom two quintiles, so the positive contribution to the endowments effects is only for the top three quintiles.
6. Endowments in terms of household size, percentage of children, gender of the household head, age of the household head, marital status, and health shocks do not explain significant changes in consumption expenditure between 2005 and 2010. This is not surprising given that levels of these endowments have changed very little or not at all over the period.
7. The data does not allow one to discern whether households remained in the same quintile in 2005 and 2010 or not. However, the results on climate shocks indicate that either (1) households that were not in the bottom quintile in 2005 had fallen into it by 2010 partly as a result of the severity of climate shocks they experienced, and/or (2) households that were already in the bottom quintile in 2005 saw their consumption decrease further by 2010 partly for the same reason.
8. Unfortunately, comparable data on prices at the community level were not collected in 2001 and 2012.
9. A study by Minten (1999) demonstrates the importance of distance to a road and "soft" infrastructure to improve competition among traders over the quality of a road for improving market integration within Madagascar, but the study is dated.
10. Unfortunately, there was no community survey in 2012 to assess whether conditions had changed.

REFERENCES

Auffret, Philippe. 2014. "Madagascar: Three Years into Crisis." Social Protection and Labor Discussion Paper 1416, World Bank, Washington, DC.

Belghith, N., T. Osborne, and P. Randriankolona. 2016. "Madagascar Poverty and Inequality Update: Recent Trends in Welfare, Employment, and Vulnerability." Washington, DC: World Bank.

Bi, C., and T. Osborne. 2016. "Transactions Costs, Poverty, and Low Productivity Traps: Evidence from Microenterprises in Madagascar." Washington, DC: World Bank.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8 (4): 436–55.

Dang, H., P. Lanjouw, J. Luoto, and D. McKenzie. 2014. "Using Repeated Cross-Sections to Explore Movements into and out of Poverty." *Journal of Development Economics*. 107: 112–28.

David-Benz, Hélène. 2011. "A Madagascar: les prix du riz flambent, sans rapport avec le marché international." Paris: Cirad, UMR Moisa.

David-Benz, Hélène, and Frederic Lancon. 2013. "Transmission des prix internationaux du riz sur les marches africains: le long terme, la crise de 2008 . . . et maintenant?" Presented at Third Africa Rice Congress, Cirad, Paris.

Deaton and Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis," unpublished.

DiNardo, Fortin, and Lemieux. 1996. "Unconditional Quantile Regressions," *Econometrica* 77 (3) : 953–973.

EPM (Enquête Périodique auprès des Ménages). 2005, 2010. INSTAT, Government of Madagascar.

FAO (Food and Agriculture Organization of the United Nations). 2010. *Mission d'évaluation de la securité alimentaire a Madagascar.* Rome: FAO.

Firpo, S., N. Fortin, and T. Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77 (3): 953–73.

Fortin, N., T. Lemieux, and S. Firpo. 2010. *Decomposition Methods in Economics.* Vancouver: University of British Columbia. http://econ.sites.olt.ubc.ca/files/2013/05/pdf_paper_thomas-lemieux-decomposition-methods-economics.pdf.

INSTAT (National Institute of Statistics, Madagascar). 2014. *Enquête national sur le suivi des objectifs du millénaire pour le développement (ENSOMD) à Madagascar.* Antananarivo.

Ivanic, M. and W. Martin. 2014. http://documents.worldbank.org/curated/en/106581468325435880/Short-and-long-run-impacts-of-food-price-changes-on-poverty.

Koenker. 2005. *Quantile Regression.* Cambridge University Press, Cambridge, UK.

Koenker and Bassett. 1978. "Regression Quantiles," *Econometrica.* 46(1): 33–50.

Machado and Mata. 2005. "Counterfactual decomposition of changes in wage distributions using quantile regression," *Applied Econometrics* 20(4): 445–465.

Moser, C., C. Barrett, and B. Minten. 2009. "Spatial Integration at Multiple Scales: Rice Markets in Madagascar." *Agricultural Economics* 40: 281–94.

Oaxaca, Ronald L. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 693–709.

Randrianarisoa, J.-C. 2003. "Analyse spatiale de la production rizicole malgache" in *Agriculture, pauvreté rurale et politiques économiques à Madagascar.* Washington, DC: USAID, Cornell University, INSTAT, and FOFIFA (Malagasy for Centre Nationale pour la Recherche Appliquée au Développent Rurale, edited by B. Minten, J.-C. Randrianarisoa, and L. Randrianarison. Ithaca, NY: Cornell Food and Nutrition Policy Program.

World Bank. 2014. *Face of Poverty in Madagascar: Poverty, Gender, and Inequality Assessment*. Washington, DC: World Bank, PREM Africa.

World Bank. 2015. *Madagascar: Systematic Country Diagnostic*. Washington DC: World Bank.


THB
AccèsBanque
AccèsBanque
MoneyGram
Bijouterie
08:24
TELL
612/CP/ABE

CHAPTER 3

Flexible Poverty Profiling and Welfare Prediction in Madagascar

Linden McBride*
Theresa Osborne
June 2016
*Cornell University
This paper was funded by the World Bank and carried out under the direction and supervision of the Africa Poverty Global Practice. Christopher Barrett (Cornell University) helped supervise the work. The authors are grateful to Kathleen Beegle and Dominique van de Walle for peer review comments.

Introduction and Key Findings

In Madagascar, where 70.7 percent of the population is below the national poverty line, an understanding of the characteristics and conditions that best predict differing levels of deprivation cannot only provide a profile of the poorest groups in the country but it can also be one fundamental step in the targeting of interventions and the design of effective policy responses. This paper applies regression tree and regression forest analysis—flexible, data-driven methods for the construction of prediction models—to identify the observable household and community-level characteristics that are the most powerful predictors of households' consumption levels.[1] These "machine learning" approaches take a different approach to conventional poverty profiling, which describes data according to arbitrarily chosen differentiating factors or inflexible (parametric) econometric estimates. By using out-of-sample predictive performance as the criterion for choosing best estimates, the approach used here develops conditional, data-driven profiles of who is poor and how poor they are. This results in a nested set of the most important factors that predict welfare, as represented by per capita consumption expenditures, thus providing a more nuanced guide for targeting of programs.[2] Like other poverty profiles, the analysis provides only correlations and is thus an insufficient basis for predicting the impacts of policies; however, the ranking of the most predictive variables is nonetheless suggestive of both further analysis and policy priorities.

This analysis utilizes the nationally representative 2010 Madagascar Enquête Périodique auprès des Ménages (EPM), the latest available household consumption survey for which community-level variables are available. The application of regression tree (RT) and random forest (RF) analyses to these data allows us to identify several clusters of variables that are highly predictive of household per capita consumption and use these to group households by similar characteristics and consumption levels. The two methods give similar but not identical results. Combining insights of each, of the many available household, regional, and community-level variables which one might expect to be correlated with poverty, those that are most predictive of increasing severity of poverty are the following, in order of importance: (1) living in a community with levels of electrification at less than 27 percent of households; (2) having a non-university-educated head of household; (3) having an illiterate head of household; (4) living in greater remoteness from the nearest major urban center, a variable which predicts welfare better than other measures of urban attributes or access to services; (5) achieving lower prices for paddy rice, and/or other indicators of performance of rice markets; and (6) having lower livestock holdings.

These results indicate that having a university education makes it very likely to have higher incomes in urban areas, as would be expected, but apart from this distinction and illiteracy, differing levels of educational attainment are not important predictors of welfare. For agricultural households, analyzed separately, the key predictive variables in order of importance are the following: (1) less cultivated land, (2) greater remoteness from the nearest major urban center, (3) having low or modest levels of electrification in the community, (4) getting higher percentage of one's revenues from agriculture, and (5) having a lower price of paddy rice.

While the importance of certain variables identified may not be surprising, using these methods one is able to sort through a long list of variables that might have otherwise been expected to play as meaningful a role in predicting levels of poverty (or consumption). For example, variables such as the household's being female headed, living in the capital, and the various regional variables do not play an important role in explaining the variation in consumption in the 2010 EPM. Thus, the specific ranking of the most predictive variables would be difficult to ascertain a priori.

Data and Methods

The methods used for the analysis include a recursive partitioning algorithm known as classification and regression tree (CART), as well as an extension of this algorithm, called random forests (RF). (See annex 3A for details.) These techniques permit a fully flexible, data-driven approach to determining the profiling parameters more useful for sorting households into increasingly homogeneous groups in terms of their characteristics, circumstances, and consumption levels. These methods offer several advantages over traditional approaches to poverty profiling. In particular, where the true underlying model is unknown, these methods are more flexible and can be more informative than parametric models that require several possibly faulty assumptions. In addition, they offer higher out-of-sample predictive accuracy than traditional methods (Breiman 2001). This feature is important because the population in which we are interested is rarely the precise one observed in the available data. In the case of Madagascar, the most recent available data (with a complete set of variables) are a sample from 2010, whereas the sample of interest is the population of all households in 2016 or later. Under the assumption that the same data-generating process underlies both the sample we have available and the population in which we may be interested, we seek a method that is most accurate for that population and not just accurate in the sample we happen to have available. Therefore, the "out-of-sample" predictive ability of this approach is both temporal and spatial. Because we want to use past survey data to try to profile the poor in a larger population and in subsequent periods, such as for the purposes of targeting interventions, out-of-sample predictive performance matters a great deal.[3]

The RT version of CART operates by recursively partitioning the data into consumption groups by variables, their threshold values, and consumption thresholds that provide the best prediction, and then into subgroups, which successively reduce error in predicting consumption levels. It essentially runs a horse race among all of the potentially predictive variables to rank their explanatory and predictive power, while splitting the population into more-poor or less-poor groups by characteristics and context, allowing the data to tell the analyst which variables are more or less important in profiling and targeting the poor.

The random forest (RF) version of the algorithm provides several innovations over and above the RT: instead of building just one regression tree, it builds hundreds (a forest), over randomly selected subsets of the data, and then averages across all of these trees in what is called "bootstrap aggregation" (or "bagging") for a final prediction (Breiman 2001).[4] With this innovation, which allows for averaging across low bias, de-correlated trees, RFs produce low-bias, low-variance predictions that are highly accurate out of sample (Breiman 2001). In addition to identifying key variables that allow for more nuanced profiling of the poor than is feasible using traditional descriptive and econometric methods, this technique allows one to rank their importance and observe how their partial correlations vary and co-vary with the outcome variable of interest. How important a role a given variable plays in predicting consumption levels is assessed by the extent to which its exclusion from the predictive model increases the out-of-sample (squared) prediction error.[5] In addition, the relationship between a single variable and consumption can be plotted by incrementally changing the variable over its range (while holding all other variables at their means) to see how the predicted response changes (Hastie, Tibshirani, and Friedman 2009). Such plots showing large jumps in the predicted response due to an incremental change in the variable of interest could suggest areas where thresholds or other nonlinearities may lie. This can reveal consumption patterns that bifurcate around values for particular household or community characteristics. While RFs offer a more robust and lower variance prediction than RTs, their results are somewhat more difficult to interpret visually.

A comprehensive set of observable household characteristics, assets and community level data are included in the analyses that follow. First are household characteristics, circumstances, and assets, including (1) the household size and dependency ratio; (2) the age, education level, sex, employment status, and marital status of the household head; (3) the number of primary-school-educated individuals in the household; (4) the ownership of productive agricultural assets such as plows, carts, harrows, and manual agricultural equipment; (5) tropical livestock units (TLUs) owned;[6] (6) the amount of land cultivated in *ares* by the household,[7] including ownership of a nonagricultural enterprise; (7) the percent of household revenue from various sources (fishing, nonagricultural enterprise, agriculture, and livestock); (8) whether the household is an agricultural household; (9) whether the household is a net consumer or net producer of paddy rice and of dehulled rice;[8] (10) whether the household lives in a rural or urban area; and (11) what climate, economic, health, security, and other shocks the household has been exposed to in the past year. Also included are a variety of community-level variables, such as (a) the level of local security; (b) the mean and standard deviation of the price and availability of white, imported, and paddy rice and inorganic fertilizer inputs (urea and nitrogen, phosphorus, potassium fertilizer (NPK) over the seasons in the local community;[9] and (c) regional dummy variables. Finally, to capture the remoteness of each community, several variables are included: distance in hours to the nearest market, health center, school, public transportation, and location to purchase agricultural inputs; and distance in kilometers to the nearest urban center and the cost of transporting 50 kilograms of rice to the nearest urban center during the wet and dry seasons. Summary statistics for all of the above variables are provided in annex table 3A.1.

Results

CART RESULTS

The distribution of log per capita consumption (in 2001 deflated ariary per person) is presented in figure 3.1, where we can see that the majority of households are consuming below the poverty line (the vertical red line), the mean log per capita consumption is 12.03, the median is 11.97, and the poverty line is 12.17. Figures 3.2 through 3.4 present the results of the RT analyses. Figure 3.2 includes the full sample; figure 3.3 includes the full sample but excludes from the search algorithm the demographic variables of household size and the dependency ratio for reasons that are detailed below; figure 3.4 includes only agricultural households and excludes demographic variables.

In the RT presented in figure 3.2, the oblong circle at each node contains the conditional mean logged per capita consumption and the percentage of the total sample that meets the condition(s) displayed at this and any preceding nodes. In the case of the very top of the tree, the conditional mean is the sample mean, a logged per capita consumption of 12.0; 100 percent of the sample is found there. Displayed on each branch descending from the first node is the first most predictive logical condition

FIGURE 3.1: Log per Capita Consumption Distribution

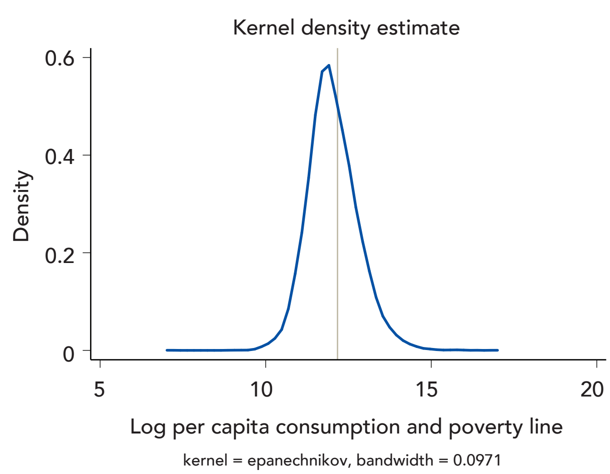


FIGURE 3.2: Regression Tree of Log per Capita Household Consumption on Household and Community Level Variables, 2010 EPM (n = 12,460)

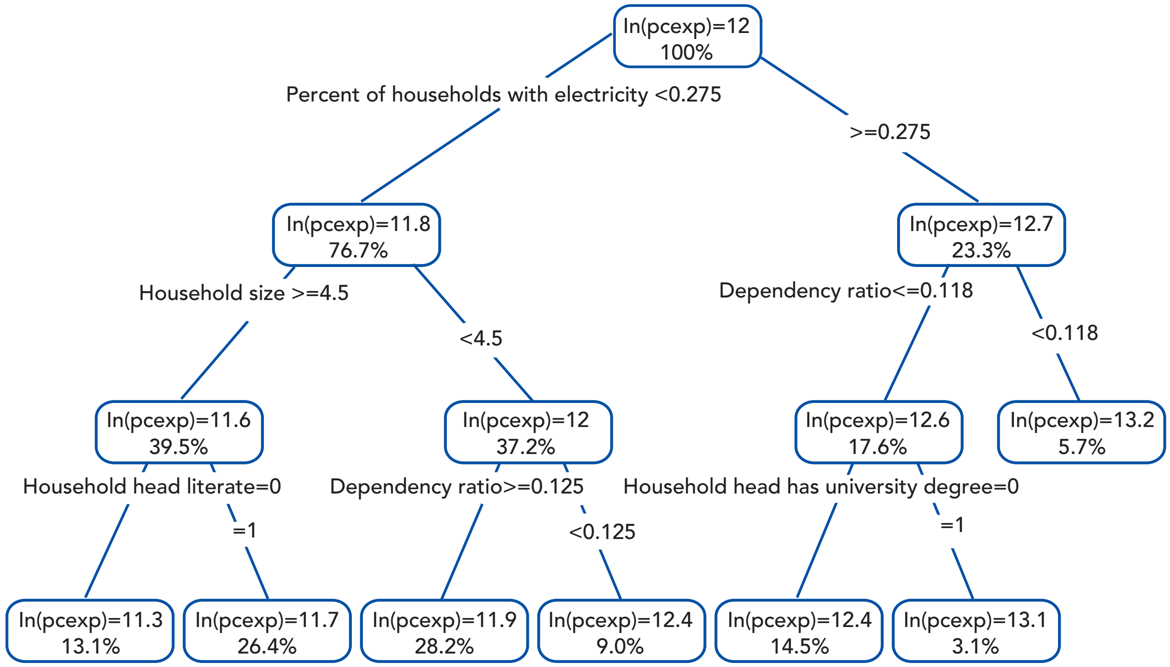


FIGURE 3.3: Regression Tree of Log per Capita Household Consumption on Household and Community Level Variables (Demographic Variables Excluded), 2010 EPM (n = 12,460)

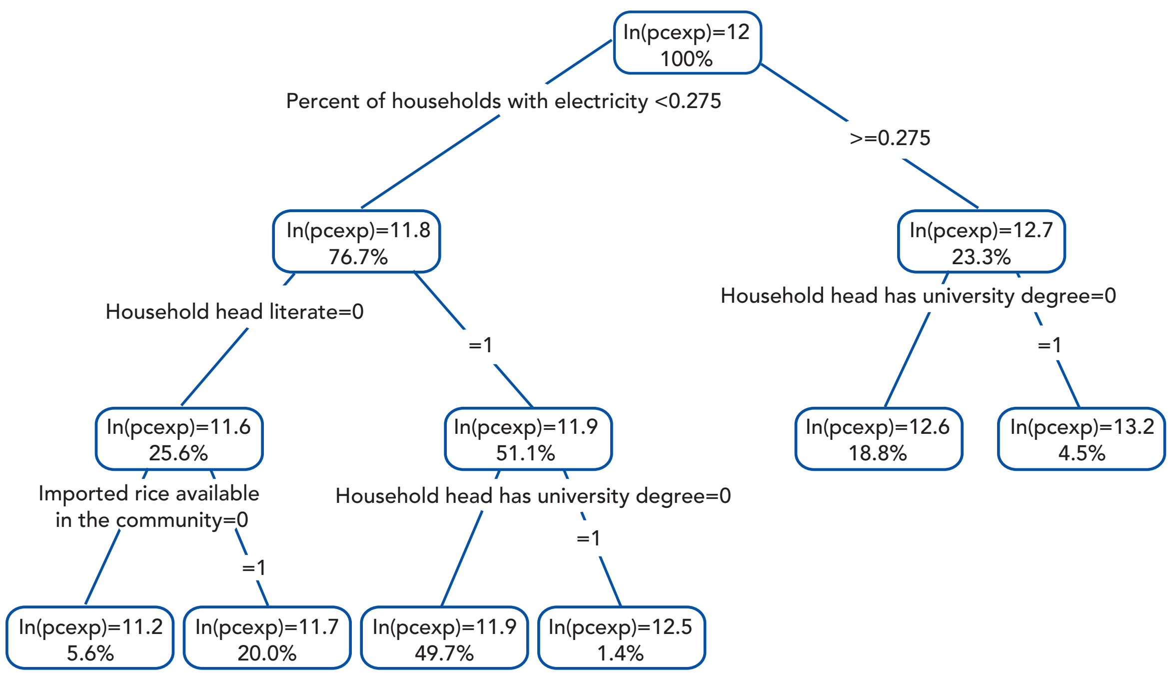
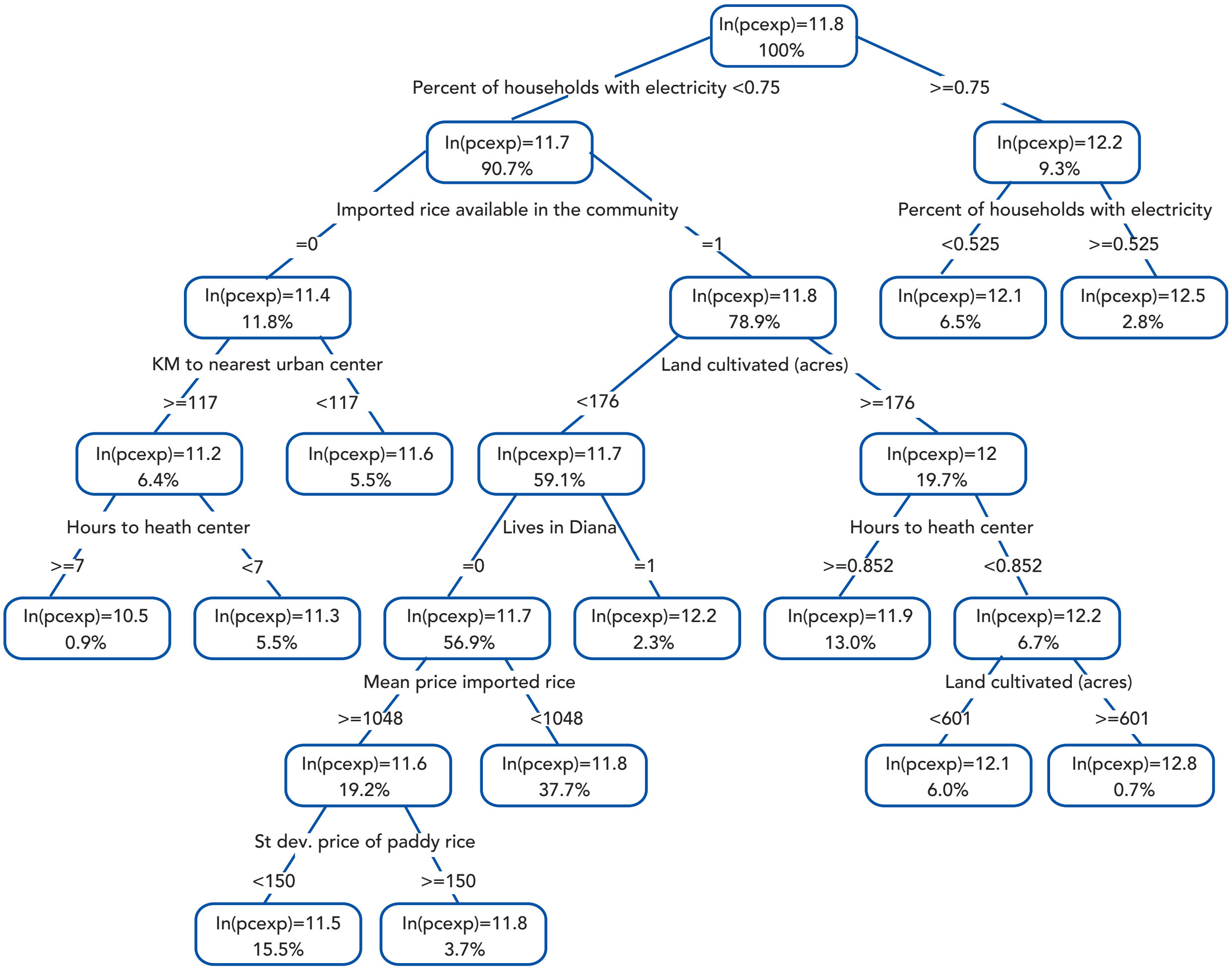


FIGURE 3.4: Regression Tree of Logged per Capita Household Expenditures on Household and Community Level Variables, 2010 EPM Agricultural Households (n = 8,145)



identified by the algorithm. If a given household's characteristics meet the logical condition presented in the left branch (in the case of figure 3.2, this first condition is whether the household is located in a community where fewer than 27.5 percent of households have electricity), then one continues down the left branch of this first node. If the household meets the condition in the opposing branch (or equivalently, fails to meet the condition in the left branch), then one continues down the right branch. As one moves down the tree, the conditional mean of consumption and the percent of households that remain in that branch (meeting the preceding conditions) are each presented at each node. Likewise, the terminal node for each branch is the mean predicted per capita expenditure for households meeting all the conditions in each of the preceding nodes in that branch of the tree. All variables listed in annex table 3A.1 are utilized, and therefore those that appear in the RT explain more variation in consumption than do those that do not appear in the tree.

From the CART analysis shown in figure 3.2, we see that households living in communities without much electricity, households with larger household size and/or larger dependency ratios, households living in communities where imported rice is not available, and heads of household with low education (illiterate) are at the lower end of the per capita consumption distribution, whereas households living in areas with more electricity, with lower dependency ratios, and headed by individuals with university-level education have higher consumption.

The electricity variable may to some extent capture differences between larger and smaller urban and increasingly rural households that are not sufficiently absorbed by the urban-rural variable and regional dummies, which are also included. In fact, the average rural household lives in a community where 4 percent of households have electricity, while the average urban household lives in a community where 35 percent of households have electricity. Moreover, the average household in the capital lives in a community where 77 percent of households have electricity, while the average household in the rest of the country lives in a community where only 18 percent of households have electricity. One cannot infer causality in the sense that access to electricity directly raises incomes and therefore consumption levels. This result could be due to the increased economic activity or wealth in the community. In this case, these community attributes would cause greater income-generating opportunities, but the inference that electricity itself causes higher income could be spurious. Alternatively, community-level electrification may proxy for the households' unobserved level of wealth.[10] We examine the likelihood of these interpretations in some of the following material.

The variables that appear in the tree in figure 3.2 allow the algorithm to capture more variance in the dependent variable, consumption, than do the variables that do not appear. Any variable that did not appear in this tree failed to improve the sum of squared prediction error of the model by a minimum of 0.007. In this respect, it may provide just as much insight to consider the variables that the algorithm does not select as those that it does, as we build a differentiated profile of poor households in Madagascar. In figure 3.2 we do not see several variables we might expect, including female-headed households, land ownership, whether the household lives in a rural or urban environment, or any information about households' livelihoods. We will explore the correlates of several of these variables further to understand whether, for example, the rural-urban and livelihoods information is being picked up by the electricity and education variables.

Interpretation of the results in figure 3.2 is complicated by the fact that household size and dependency ratio variables capture both welfare and measurement issues. Although there is a real relationship between dependency and welfare, whenever one uses per capita consumption as the welfare indicator, one will overstate the relative impact of household size on household members' welfare, as this indicator fails to adjust for age-specific consumption needs and household-level economies of scale through an adult equivalence scale (see, for example, Deaton and Zaidi 1999; Deaton 1997). Because household composition has large effects on per capita consumption by construction, for the rest of the analysis, these demographic variables—household size and the dependency ratio—are omitted.

Figure 3.3 displays the results with basic demographic variables omitted. Absent household size and the dependency ratio, access to electricity remains a key explanatory variable, and education emerges as even more important—in particular, whether the household head is illiterate or has a university education. Moreover, the availability of imported rice in the community and whether or not the household is in the Diana region—the northern-most region of Madagascar where many households rely on fishing, forest products, and agriculture for their livelihoods—bifurcates the households having among the lowest consumption levels. Among households with slightly greater consumption levels (just below the consumption poverty line but above the most destitute households in the sample), the remoteness variable, "hours to nearest health center" bifurcates households, placing those traveling for longer than 0.85 hours to a health center in a lower consumption branch. Other indicators of access to services and remoteness are less important predictors than this one.

Next, we perform the same analysis over the subset of 8,145 households in which the head of household reported agricultural work as his or her primary income-generating activity (with demographic variables again excluded). The resulting regression tree is presented in figure 3.4, where we can see that electricity again serves as the first splitting factor as well as the variable that bifurcates the households with the highest consumption levels (log per capita consumption of 12.5) from those with slightly lower consumption (log per capita consumption of 12.1).

Following the leftmost branch of figure 3.4 from the root node, we see that the agricultural households with the lowest consumption levels (log per capita consumption of 10.5) are found in communities that lack availability of imported rice, are far from the nearest urban centers, and are far from the nearest health center: the lowest-consumption agricultural households are those living in the remotest areas. Among households in the middle of the expenditure distribution, having larger land holdings and living in less remote areas are associated with slightly higher consumption levels. Meanwhile, having smaller land holdings, household residence in any region besides Diana, and facing higher imported rice prices and lower standard deviation of paddy rice prices are associated with slightly lower consumption. Note that, in comparison to the full sample, education variables (such as literacy and having completed university) are not differentiating factors for the consumption levels among agricultural households. Education primarily differentiates consumption levels between agricultural and nonagricultural households.

While these results provide clear profiling information, the implications for food and policy and public investments in infrastructure are less clear. To aid our interpretation, table 3.1 shows that the price, standard deviation, and the local availability of rice differ significantly by type. White and imported rice are both more expensive and more available than is paddy rice in local markets. Although imported rice is slightly more expensive than "white" rice, it has the lowest standard deviation of all rice prices across the seasons, and thus could be less expensive in certain seasons. Imported rice is available in 80 percent of communities, but is not quite as ubiquitous as the slightly less expensive white rice, available in 98 percent.

One would expect rice prices to affect rice producers and net rice consumers differently. Households in 2010 generally sold paddy rice and purchased dehulled rice. We classify households as net consumers or net producers of each type of rice based on whether they bought or produced more by weight for each type of rice. As shown in table 3.2, 63.7 percent of households are net producers of paddy rice, that is, they produce more paddy rice than they buy. However, only 0.4 percent are net producers of dehulled rice. Meanwhile, only 2.0 percent of households are net consumers of paddy rice while 72.4 percent are net consumers of dehulled rice.[11] Overall, 4.3 percent of the population is involved in neither production nor consumption of either type of rice. The fact that the local nonavailability of imported rice bifurcates those households at the lower end of the consumption distribution suggests that the variable may proxy for local community purchasing power, that is, the availability of this rice is lower where households are less able to afford it. Or it may proxy for market integration and/or the preferences of rice consumers in some remote areas. Moreover, where it is available, but at a higher price than 1,048 ariary per kilogram, only households with extremely low landholdings are poorer. Such households would tend to be net consumers of rice, and a lower imported rice price would tend to push down other local prices (figure 3.4). Further insight on these questions is gained from the RF analysis that follows.

RANDOM FOREST RESULTS

To pin down the importance ranking of predictive variables, we next ran both the full and agricultural-household-only datasets through the RF algorithm. Because RFs are more robust than single-regression trees, they can help confirm and extend several of the insights

TABLE 3.1: Mean and Standard Deviation of Seasonal Prices (per Kilogram) and Local Availability of Rice by Type, at the Community Level

Type	Mean price across seasons	SD of prices across seasons	Availability of rice type in local community
White rice	974.98	169.73	97.5%
Paddy rice	730.00	168.39	49.8%
Imported rice	997.60	99.27	80.3%

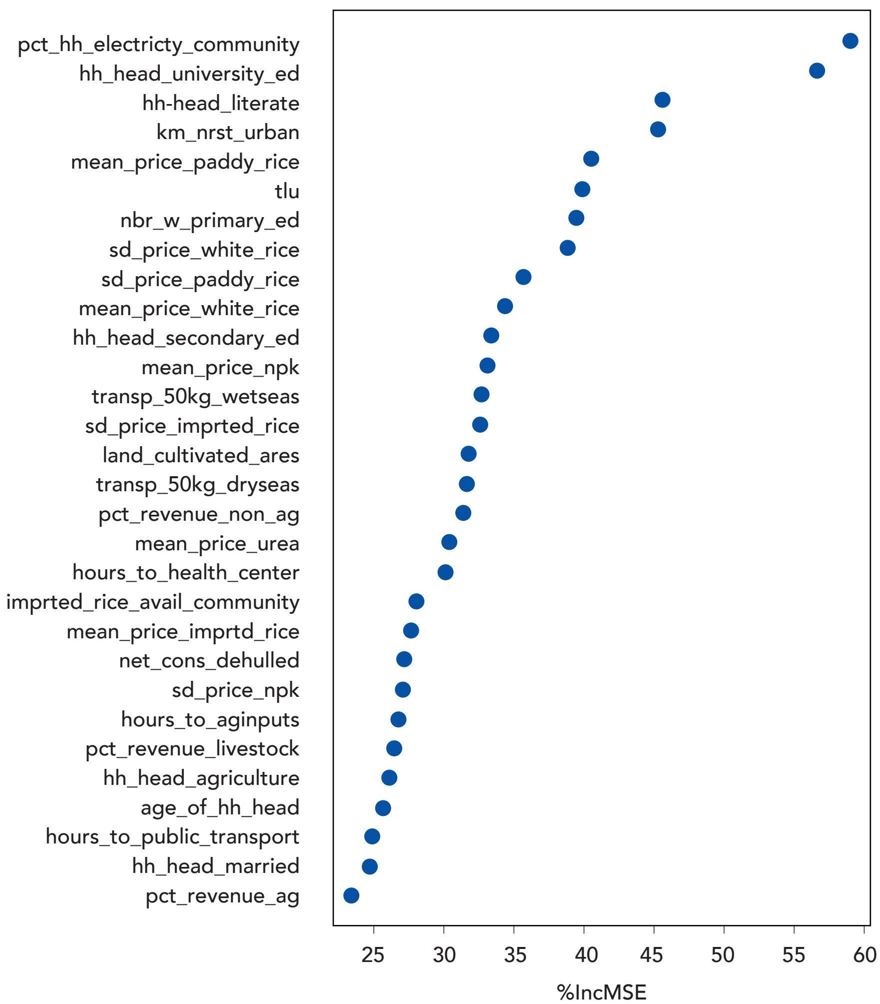
TABLE 3.2: Consumer or Producer Status by Type of Rice (Population Weighted)

	Net producer	Net consumer	Neither net producer nor net consumer
Paddy rice	63.7%	2.0%	34.3%
Dehulled rice	0.4%	72.4%	27.2%

from the RTs. However, RFs are more difficult to interpret than RTs as they cannot offer a single branching figure displaying the conditional relationships and interactions among the variables. Therefore, we report two types of output from the RF algorithm: variable importance plots (figures 3.5 and 3.7) and partial dependence plots (figures 3.6 and 3.8). The variable importance measures how great a role a given variable plays in reducing the error of the out-of-sample prediction across the forest, while partial dependence plots display the effects of variables of interest on the forest's prediction of consumption.

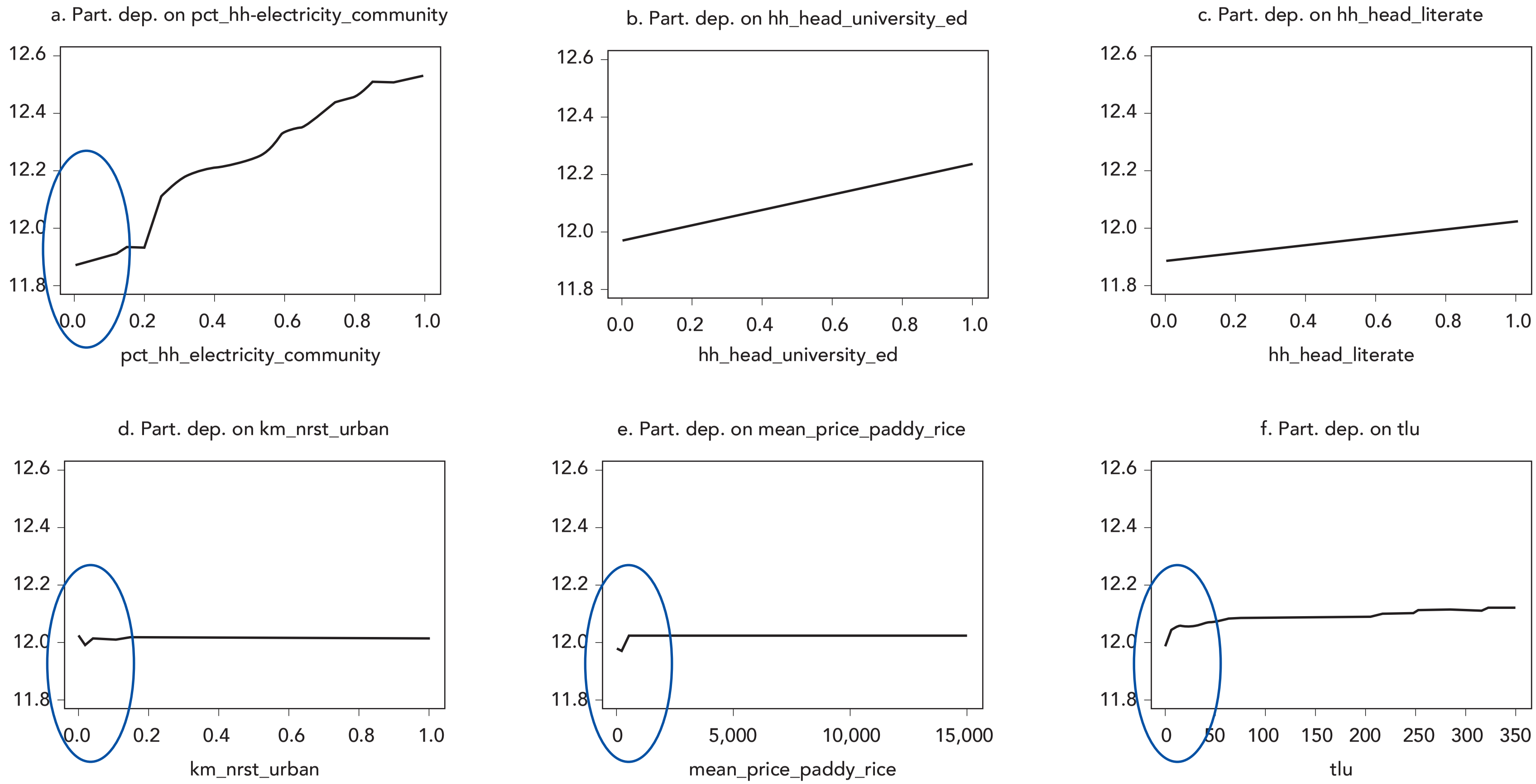
In the RF variable importance plot for the full sample (figure 3.5) we see, as anticipated by the single RT analysis reported in figure 3.3, that variables such as the percent of households with electricity in the community and whether the head of household has a university degree or is literate play a large role in reducing the out-of-sample prediction error. In addition, the remoteness variable, "kilometers to the nearest urban center," the mean price of paddy rice in the community, the TLU holdings of the household, and the number of primary school educated members in the household play a more substantial role than the other variables in reducing the out-of-sample prediction error.

FIGURE 3.5: Variable Importance Plot, 2010 EPM (n = 12,460)



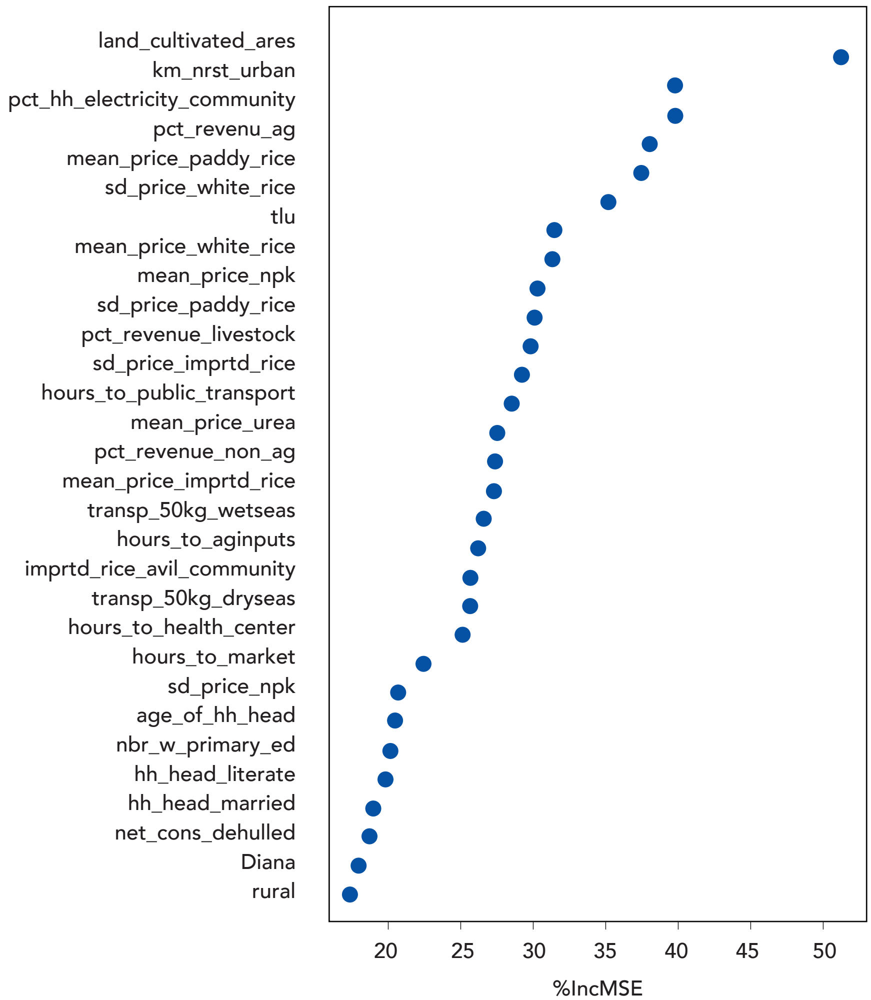
Note: The x-axis of this dot plot is percent increase in mean-squared error (MSE). %IncMSE is the percent that MSE of predicted logged per capita income increases due to the perturbance of this variable. *pct_hh_electricity_community* is the percent of households in the community with electricity; *hh_head_university* is a binary variable indicating whether or not the household head has completed university; *hh_head_literate* is a binary variable indicating whether or not the head of household is literate; *sd_price_white_rice* is the standard deviation in the price of white rice across the seasons as available in the local community; *nbr_w_primary_ed* is the number of members of the household with a primary school education. EPM = Enquête Périodique auprès des Ménages.

FIGURE 3.6: Partial Dependence Plots, 2010 Log per Capita Consumption (y Axis) on Key Variables (n = 12,460)



Note: pct_hh_electricity_community is the percent of households in the community with electricity; *hh_head_university* is a binary indicating whether or not the household head has completed university; *sd_price_white_rice* is the standard deviation in the price of white rice across the seasons as available in the local community; *mean_price_white_rice* is the mean price of white rice across the seasons as available in the local community; *hh_head_literate* is a binary indicating whether or not the head of household is literate; *nbr_w_primary_ed* is the number of members of the household with a primary school education. Part_dep=partial dependence.

FIGURE 3.7: Variable Importance Plot, 2010 EPM Agricultural Households (n = 8145)

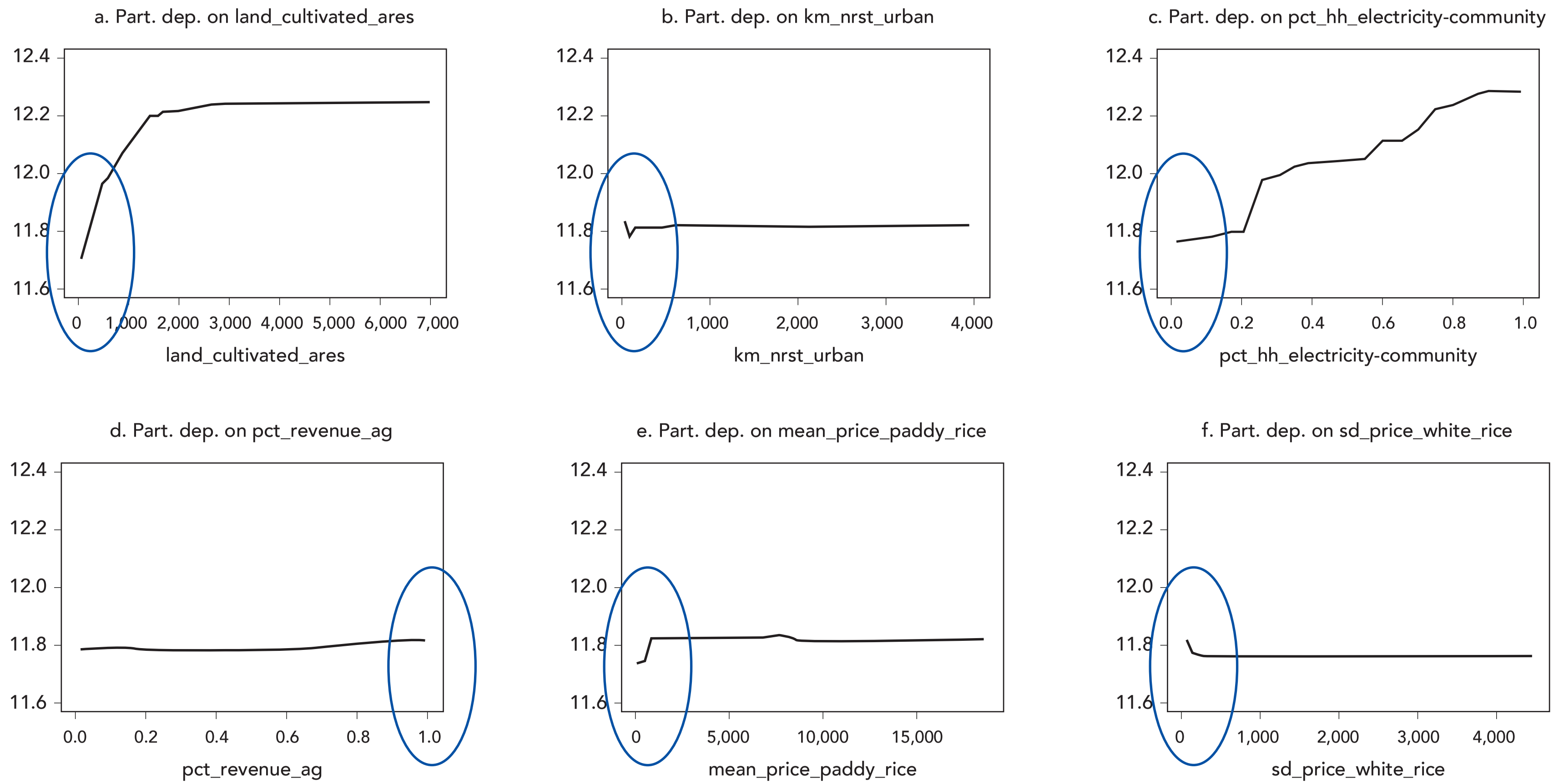


Note: The x-axis of this dot plot is percent increase in MSE. %IncMSE is the percent that MSE of predicted logged per capita income increases due to the perturbance of this variable. *land_cultivated_ares* is the land area cultivated by the household in the local measurement unit of *ares*; *pct_hh_electricity_community* is the percent of households in the community with electricity; *mean_price_paddy_rice* is the mean price of paddy rice across the seasons as available in the local community; *remoteness* is an index capturing how remote the community is in terms of access to services; *sd_price_white_rice* is the standard deviation in the price of white rice across the seasons as available in the local community. EPM = Enquête Périodique auprès des Ménages.

In considering variable importance for prediction of per capita expenditures among agricultural households only (figure 3.7) we see a somewhat different ranking of variables by predictive importance. Notably, land area cultivated (*ares*) and the percentage of revenue from agricultural activities emerge as more important in this subset of the data. Consumption is increasing with the area of cultivated land, as would be expected in a context with such small farm sizes, and is decreasing in the percentage of revenues from agriculture. In addition, as with the full sample, we see that distance to the nearest urban center, electrification in the community, the mean price of paddy rice, and TLU holdings are important predictors in correctly predicting where an agricultural household will lie on the expenditure distribution.

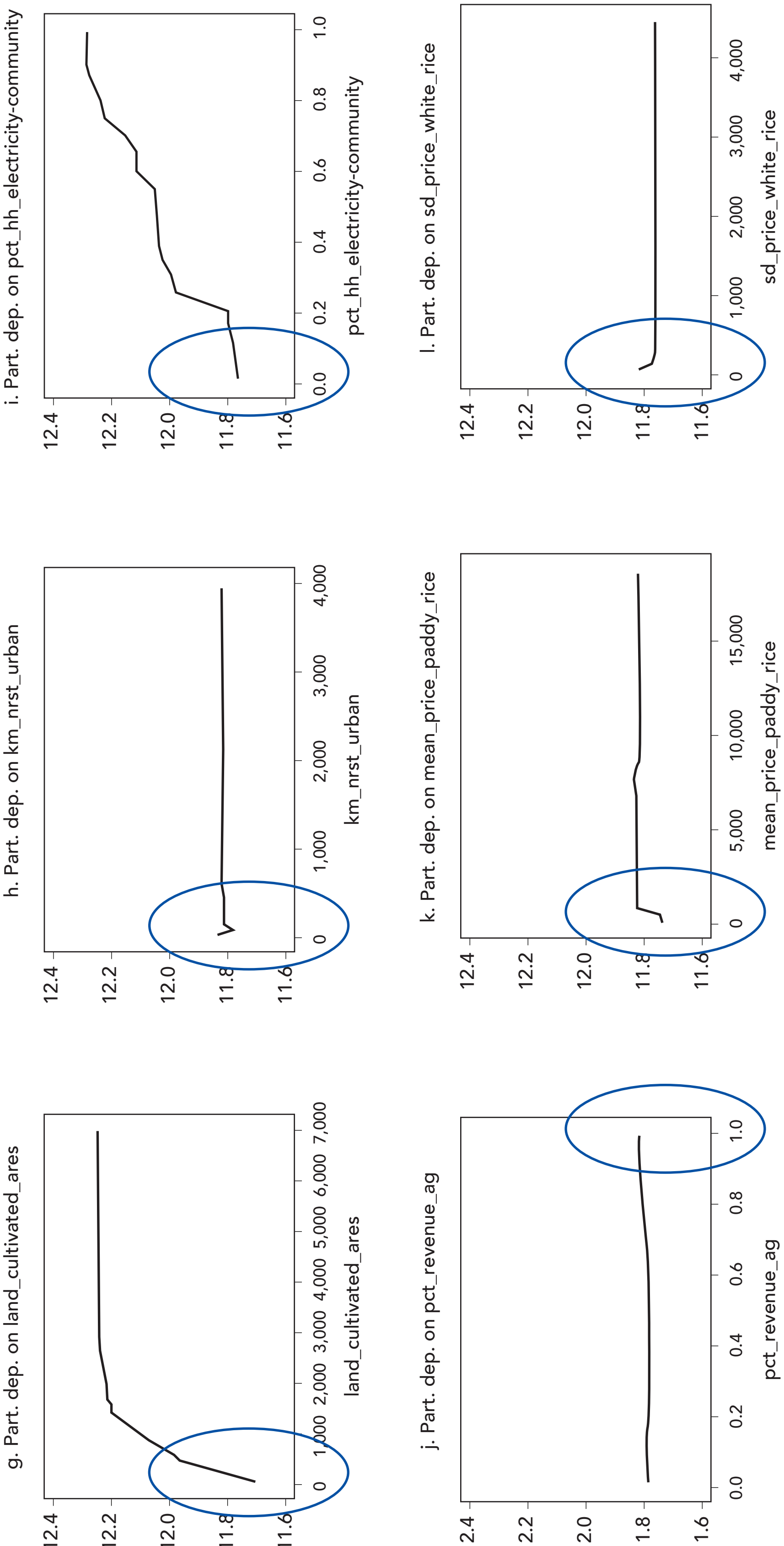
The RF results for both the full and agricultural household samples yield a different ranking of key variables related to outcomes in local rice markets and thus a different interpretation of results. Both analyses underscore the performance of markets for the country's staple food and dominant crop for poverty reduction. Because the different indicators in rice markets are related and may have nonlinear effects, however, the ranking is sensitive

FIGURE 3.8: Partial Dependence Plots, 2010 Log per Capita Consumption (y Axis) on Key Variables, Agricultural Households (n = 8,145) (Relevant Range Circled)



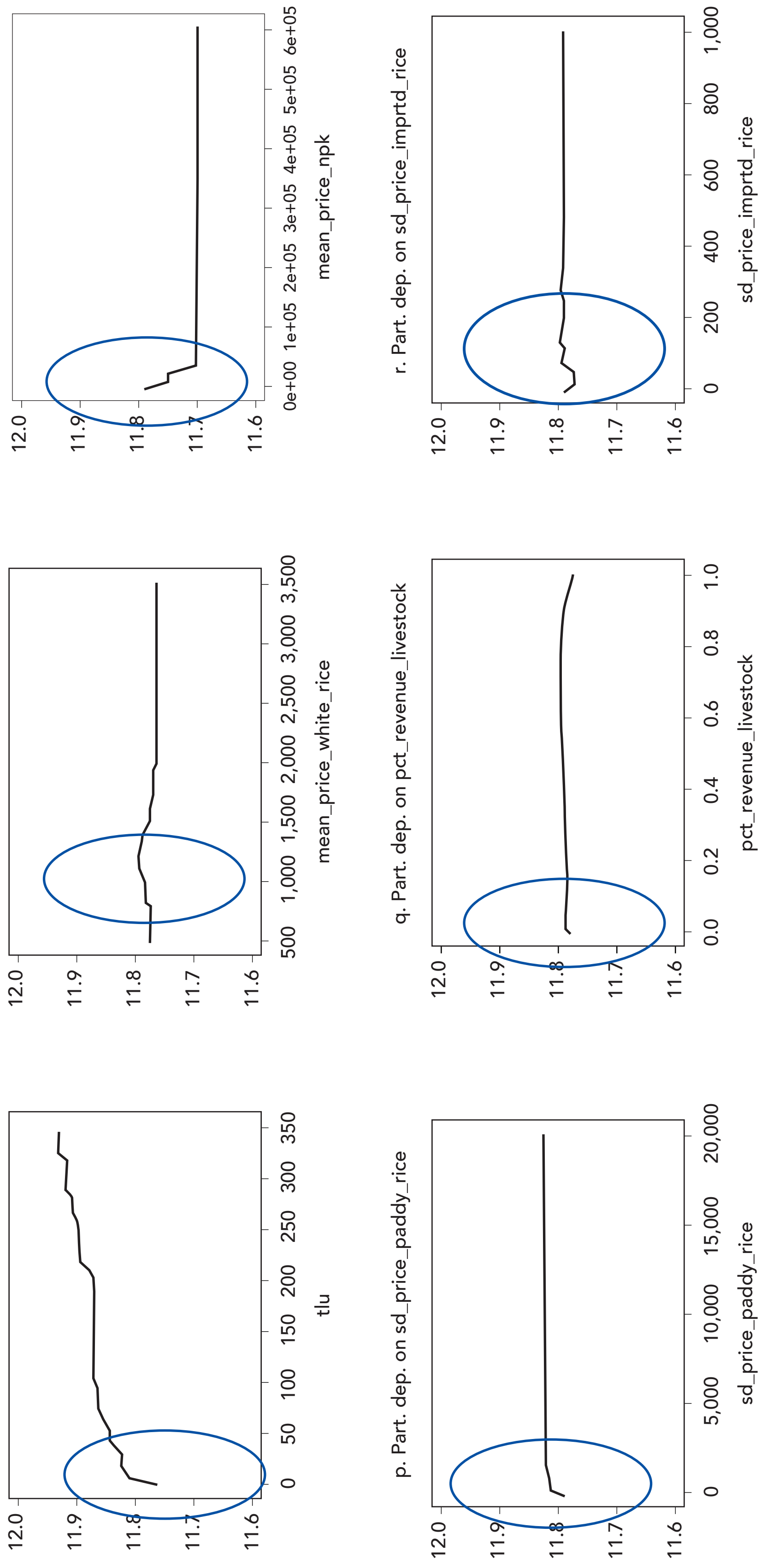
(continued)

FIGURE 3.8: *Continued*



(continued)

FIGURE 3.8: *Continued*



Note: part. Dep. = partial dependence.

to the precise methodology used. Given that RF provides a more accurate out-of-sample prediction and better illustrates possible nonlinearities, we derive our interpretation of results on rice markets from this methodology. The RF results suggest that the availability and price of imported rice is not as important a determinant of welfare as in the RT results. Rather, the mean price of paddy rice and standard deviations of the price for white rice and paddy rice ranking higher in importance. The availability of imported rice falls to 20th place in the ranking of variables, below the price of transport, the price of urea fertilizer, and hours to the nearest health center.

To better observe the role of each of these important predictors, figures 3.6 and 3.8 display their partial dependence plots, that is, the relationships between the variables on the horizontal axis and predicted log per capita consumption (on the vertical axis), holding all other variables at their mean values. The “rug plots” at the bottom of each plot indicate the data density, with circles to highlight the relevant ranges. In figure 3.6, where partial dependence plots are reported for the full sample analysis, one observes a steep relationship between the percentage number of households with electricity in the community and welfare. For the other variables, the marginal effects (represented by the slopes of the curves) are not as pronounced. Rather, for the binary education variables (household head has a university degree and household head is literate) we see slopes indicative of a positive marginal effect across the mean range of the household consumption distribution. For continuous variables, such as kilometers to the nearest urban center, the mean price of paddy rice, and TLU holdings, we see clear slopes where the bulk of the data lie, as indicated by the blue circles.

Figure 3.8 reports partial dependence plots for the most important variables (those reported in figure 3.7) for agricultural households only. The marginal effect of land area cultivated appears to be significant at low levels of cultivated land, where the observations are most dense. The curve flattens out at approximately 1.5 hectares (1,500 *ares*), but as shown the cultivated areas per household tend to be much lower. TLU holdings follow a similar though much more gradual trajectory as that seen for land. The distance to the nearest urban center has a negative correlation with consumption in the relevant range, as expected. Finally, the mean price of paddy rice has a positive relationship with consumption in the range where the data are available.

TABLE 3.3: Multivariate Correlates of Electrification Variable (N = 12,460, Region Dummies Included)

Dependent variable: % HH with electricity	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rural (0/1)	–0.445***	–0.439***	–0.409***	–0.406***	–0.407***	–0.386***	–0.361***
	(0.00473)	(0.00472)	(0.00474)	(0.00473)	(0.00472)	(0.00480)	(0.00443)
Owns nonagricultural enterprise		0.0570***	–0.0392***	–0.0372***	–0.0368***	–0.0370***	–0.0293***
		(0.00431)	(0.00559)	(0.00558)	(0.00556)	(0.00547)	(0.00496)
Percentage of revenue nonagriculture			0.195***	0.189***	0.188***	0.177***	0.136***
			(0.00747)	(0.00746)	(0.00745)	(0.00736)	(0.00672)
Cost of transporting 50 Kg rice, wet season				–0.00000113	–0.00000133*	–0.00000155**	–0.00000213***
				(0.000000721)	(0.000000719)	(0.000000714)	(0.000000688)
Cost of transporting 50 Kg rice, dry season				–0.00000258***	–0.00000177**	–0.000000216	0.00000148*
				(0.000000891)	(0.000000894)	(0.000000893)	(0.000000883)
Distance to nearest urban center					–0.000106***	–0.000111***	–0.0000304**
					(0.0000132)	(0.0000130)	(0.0000132)
Hours to reach nearest public transport						–0.00355***	–0.00146***
						(0.000221)	(0.000215)
Hours to reach nearest school						0.000480	0.00136*
						(0.000803)	(0.000731)
Hours to reach agricultural input supplier						–0.000672***	–0.000685***
						(0.000219)	(0.000213)
Hours to nearest health center						–0.00136***	–0.00101***
						(0.000421)	(0.000380)
Hours to nearest food market						0.00248***	0.00127***
						(0.000363)	(0.000357)
_cons	0.503***	0.478***	0.438***	0.452***	0.460***	0.475***	0.650***
	(0.00409)	(0.00446)	(0.00462)	(0.00482)	(0.00491)	(0.00495)	(0.00564)
R-squared	0.415	0.423	0.453	0.457	0.460	0.478	0.582

Note: Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To ascertain why the electrification variable plays such a large role throughout these analyses and to better understand what other things this variable may be capturing, a closer look at the correlates of this variable via multivariate regression are provided in table 3.3. Both livelihood variables and remoteness variables are considered in this examination of correlates. In table 3.3's first regression (column 1), we see that whether the household is located in a rural environment or not is statistically significant and explains 42 percent (R-squared) of the variation in the electrification variable. Subsequent regressions in this table, in which livelihood, remoteness, and regional dummy variables are progressively added, decrease the magnitude of the rural variable slightly, but do not decrease its significance. In regressions 2 and 3, livelihood variables such as ownership of and percent revenue from nonagricultural enterprise are added. When percent revenue from nonagricultural enterprise is added to the regression in column 3, we see that the ownership coefficient switches signs, suggesting that when the amount of income derived from nonagricultural income is accounted for, nonagricultural-enterprise-owning households are less likely to reside in electrified areas, but those that do live in such areas derive more of their income from nonagricultural endeavors.

Further investigation shows that while the costs of transporting 50 kilograms of rice during the dry and wet seasons (column 4) are not significantly correlated with electrification, the distance of the household from the nearest urban center is significant and negatively correlated with electrification (column 5), as we might expect. In column 6, several proxies for remoteness, used throughout the regression tree and forest analyses, are included. All but hours to the nearest school are significant. Of the significant coefficients, all are negative except hours to the nearest market. In the final column, regional dummy variables are included (a breakdown of electrification by region is shown in annex table 3A.2); although the coefficients on these regional dummies are suppressed here, all are statistically significant. With all variables included, the final regression "explains" 58 percent (adjusted R-squared) of the variation in electrification across communities. While this is a high adjusted R-squared, it still leaves much of the variation in electrification unexplained, suggesting that things we cannot observe in the data—perhaps differences in opportunity, population density and market size, the costs of delivering electricity, or political connectedness of some communities—are also driving this variable.

Discussion and Conclusion

Several clusters of variables emerge from the analyses as highly predictive of a household's falling along the high or low end of the expected per capita consumption distribution. In the full sample, the poorest households are those found in communities where fewer households have electricity, a variable that is correlated with remoteness, livelihood strategies, and regions but may also be correlated with unobserved factors such as market and educational opportunities that we cannot observe. In the full sample, household-level features such as having a literate or university-educated head of household also play a large role in separating higher and lower consumption households. The poorest households have an illiterate head of household, while the wealthiest households have a head with a university degree. Although heads of household with university degrees are not more likely to be employed (91 percent employed) than those without (95 percent employed), they are much more likely to be living in an urban environment than a rural one as well, and are more likely to be living in the capital than not. Literacy follows the same employment and environment pattern.

Yet among agricultural households, educational attainment does not appear to be an important predictor of expenditures. Rather, other productive assets such as land area and livestock holdings (below a certain threshold), the market prices of farm outputs and inputs, community-level electrification, and several proxies for remoteness (distance from nearest urban center, distance from nearest health center) appear to play a larger role. The poorest households live in the remotest areas and have low land and livestock holdings. On the list of variables that are not as predictive of consumption per capita are ownership of agricultural equipment, regional indicators, gender of household head, and marital status.

These findings provide implications for targeting and guideposts for key policy areas. In particular, they imply the targeting of interventions to reach households with especially low land holdings, in communities with productive potential but lacking electricity, and to better connect those in more remote areas. Although the issues associated with rice policies are complex and merit further investigation, these results indicate that on the whole higher producer prices aid poverty reduction.

Annex 3A. Explanation of Methods

Because the data available for the analysis of poverty in Madagascar are not ideal for obtaining sound identification for the purposes of inference, this paper takes the approach of predictive analytics. Predictive analytics differ from traditional regression analysis in several fundamental ways, and therefore offer several advantages. First, these methods target prediction of an outcome over and above parameterization of a model. Second, and relatedly, these methods make a bias for variance trade-off such that they do not produce unbiased coefficient estimates in the manner of OLS. Rather they produce highly accurate out-of-sample predictions with minimal variance (Kleinberg, Mullainathan, and Obermeyer 2015). OLS, as the best linear unbiased estimator, does not allow for such trade-offs. Third, these methods allow for nonparametric analysis with unlimited interactions and without a predefined functional form; instead, the data define the form. In this paper, the regression tree and regression forest analyses are implemented in R using packages developed by Therneau, Atkinson, and Ripley (2015) and Liaw and Wiener (2002), respectively.

TABLE 3A.1: Summary Statistics, EPM 2010 (N = 12,460, Household Survey Weights Applied)

Variable	Variable name	Mean (household weighted)	Linearized std. err.	95% conf	Interval
Household size	*hh_size*	4.76	0.03	4.70	4.81
Age of head of household	*age_of_hh_head*	41.96	0.16	41.63	42.28
Head of household literate (y/n)	*hh_head_literate*	0.73	0.00	0.72	0.74
Head of household has completed primary school	*hh_head_primar~d*	0.30	0.01	0.29	0.31
Head of household has completed secondary school	*hh_head_second~d*	0.15	0.00	0.14	0.15
Head of household has completed university	*hh_head_univer~d*	0.06	0.00	0.05	0.06
Number of households in the ea with electricity	*pct_hh_electrity_community*	0.17	0.00	0.16	0.18
Percent of revenue from fishing	*pct_revenue_fish*	0.03	0.00	0.02	0.03
Percent of revenue from nonagricultural enterprise	*pct_revenue_no~g*	0.26	0.00	0.26	0.27
Percent of revenue from agriculture	*pct_revenue_ag*	0.51	0.00	0.51	0.52
Percent of revenue from livestock	*pct_revenue_livestock*	0.11	0.00	0.10	0.11
Own agricultural cart (y/n)	*owns_ag_cart*	0.08	0.00	0.08	0.09
Own plow (y/n)	*owns_ag_plow*	0.10	0.00	0.10	0.11
Own harrow (y/n)	*owns_ag_harrow*	0.08	0.00	0.07	0.08
Owns agricultural equipment (y/n)	*owns_ag_equip*	0.77	0.00	0.76	0.78
Own nonagricultural enterprise (y/n)	*owns_non_ag_enterprise*	0.35	0.01	0.34	0.36
Number of household members with a primary school education	*nbr_w_primary_ed*	1.11	0.01	1.08	1.14
Tropical livestock units owned by household	*tlu*	1.79	0.07	1.65	1.92
Dependency ratio	*depr*	0.43	0.00	0.43	0.44
Head of household is female (y/n)	*hh_head_female*	0.20	0.00	0.19	0.21
Head of household married (y/n)	*hh_head_married*	0.75	0.00	0.74	0.76
Head of household divorced or separated (y/n)	*hh_head_div_sep*	0.10	0.00	0.09	0.11
Head of household is widowed (y/n)	*hh_head_widowed*	0.09	0.00	0.08	0.09
Head of household is employed (y/n)	*hh_head_employed*	0.95	0.00	0.95	0.96

(continued)

TABLE 3A.1: *Continued*

Variable	Variable name	Mean (household weighted)	Linearized std. err.	95% conf	Interval
Household is agricultural household	*hh_head_agriculture*	0.68	0.01	0.67	0.69
Household lives in community with bad or very bad security conditions	*bad_security*	0.33	0.01	0.32	0.34
Household living in community with average security conditions	*ok_security*	0.31	0.01	0.30	0.32
Household lives in rural area (y/n)	*rural*	0.75	0.00	0.74	0.76
Mean community price of white rice across seasons	*mean_price_white_rice*	974.98	2.85	969.40	980.56
Standard deviation of community price of white rice across seasons	*sd_price_white_rice*	169.73	4.83	160.26	179.20
Mean community price of imported rice across seasons	*mean_price_imprtd_rice*	997.60	4.47	988.84	1006.35
Standard deviation of community price of imported rice across seasons	*sd_price_imprtd_rice*	99.28	3.87	91.70	106.85
Mean community price of paddy rice across seasons	*mean_price_paddy_rice*	730.00	8.70	712.95	747.05
Standard deviation of community price of paddy rice across seasons	*sd_price_paddy_rice*	168.39	4.81	158.95	177.83
Mean community price of npk across seasons	*mean_price_npk*	5720.18	422.66	4891.70	6548.65
Standard deviation of community price of npk across seasons	*sd_price_npk*	176.40	11.74	153.38	199.41
Mean community price of urea across seasons	*mean_price_urea*	1564.03	8.01	1548.33	1579.73
Standard deviation of community price of urea across seasons	*sd_price_urea*	62.22	1.82	58.65	65.80
White rice available in community (y/n)	*white_rice_avail*	0.97	0.00	0.97	0.98
Paddy rice available in community (y/n)	*paddy_rice_avail*	0.50	0.01	0.49	0.51
Imported rice available in community (y/n)	*imprtd_rice_avail*	0.80	0.00	0.79	0.81
Npk available in community (y/n)	*npk_avail_community*	0.39	0.01	0.38	0.40
Urea available in community (y/n)	*urea_avail_comunity*	0.36	0.01	0.35	0.37
Land cultivated (ares)	*land_cultivated_ares*	101.94	2.08	97.85	106.02
Net producer of paddy rice (y/n)	*net_prod_paddy*	0.64	0.01	0.63	0.65
Net consumer of paddy rice (y/n)	*net_cons_paddy*	0.02	0.00	0.02	0.02
Net producer of dehulled rice (y/n)	*net_prod_dehulled*	0.00	0.00	0.00	0.01
Net consumer of dehulled rice (y/n)	*net_cons_dehulled*	0.72	0.01	0.71	0.73
Household lives in capital (y/n)	*capital*	0.07	0.00	0.06	0.08
Household experienced a climate shock (y/n)	*climate_shock*	0.34	0.01	0.33	0.35
Household experienced an economic shock (y/n)	*economic_shock*	0.10	0.00	0.09	0.11
Household experienced a health shock (y/n)	*health_shock*	0.06	0.00	0.05	0.06
Household experienced a security shock (y/n)	*security_shock*	0.06	0.00	0.05	0.06
Household experienced other shock (y/n)	*other_shock*	0.01	0.00	0.00	0.01
Hours from community to nearest market	*hours_to_market*	3.62	0.11	3.41	3.82
Hours from community to nearest health center	*hours_to_health_center*	2.75	0.08	2.59	2.91

(continued)

TABLE 3A.1: *Continued*

Variable	Variable name	Mean (household weighted)	Linearized std. err.	95% conf	Interval
Hours from community to location where ag inputs can be purchased	*hours_to_aginputs*	9.51	0.15	9.22	9.80
Hours from community to nearest school	*hours_to_school*	0.94	0.02	0.89	0.99
Hours from community to nearest public transportation	*hours_to_public_transp*	9.14	0.14	8.88	9.41
Kilometers from community to nearest urban center	*km_nrst_urban*	92.60	1.39	89.88	95.31
Cost of transporting 50kg of rice to nearest urban center, wet season	*transp_50kg_wetseas*	4210.15	70.77	4071.43	4348.86
Cost of transporting 50kg of rice to nearest urban center, dry season	*transp_50kg_dryseas*	3873.09	57.93	3759.54	3986.64
Household located in Analamanga (Y/N)	*Analamanga*	0.17	0.01	0.16	0.18
Household located in Vakinankaratra (Y/N)	*Vakinankaratra*	0.08	0.00	0.07	0.08
Household located in Itasy (Y/N)	*Itasy*	0.03	0.00	0.02	0.03
Household located in Bongolava (Y/N)	*Bongolava*	0.02	0.00	0.02	0.02
Household located in MatsiatraAmbony (Y/N)	*MatsiatraAmbony*	0.05	0.00	0.04	0.05
Household located in AmoroniMania (Y/N)	*AmoroniMania*	0.03	0.00	0.03	0.03
Household located in VatovavyFitovi~y (Y/N)	*VatovavyFitovi~y*	0.06	0.00	0.05	0.06
Household located in Ihorombe (Y/N)	*Ihorombe*	0.02	0.00	0.01	0.02
Household located in AtsimoAtsinanana (Y/N)	*AtsimoAtsinanana*	0.03	0.00	0.03	0.04
Household located in Atsinanana (Y/N)	*Atsinanana*	0.06	0.00	0.06	0.07
Household located in Analanjirofo (Y/N)	*Analanjirofo*	0.05	0.00	0.05	0.06
Household located in AlaotraMangoro (Y/N)	*AlaotraMangoro*	0.05	0.00	0.04	0.05
Household located in Boeny (Y/N)	*Boeny*	0.04	0.00	0.03	0.04
Household located in Sofia (Y/N)	*Sofia*	0.06	0.00	0.05	0.06
Household located in Betsiboka (Y/N)	*Betsiboka*	0.01	0.00	0.01	0.01
Household located in Melaky (Y/N)	*Melaky*	0.01	0.00	0.01	0.01
Household located in AtsimoAndrefana (Y/N)	*AtsimoAndrefana*	0.06	0.00	0.06	0.07
Household located in Androy (Y/N)	*Androy*	0.03	0.00	0.03	0.03
Household located in Anosy (Y/N)	*Anosy*	0.03	0.00	0.03	0.03
Household located in Menabe (Y/N)	*Menabe*	0.03	0.00	0.02	0.03
Household located in Diana (Y/N)	*Diana*	0.04	0.00	0.04	0.05

Std. err = standard error

TABLE 3A.2: Electrification by Region

Region	Electrification
Analamanga	48.5%
Vakinankaratra	11.0%
Itasy	13.6%
Bongolava	4.4%
Matsiatra Ambony	10.1%
Amoron'i Mania	4.9%
Vatovavy Fitovinany	4.2%
Ihorombe	5.2%
Atsimo Atsinanana	2.5%
Atsinanana	22.5%
Analanjirofo	10.0%
Alaotra Mangoro	10.4%
Boeny	23.7%
Sofia	7.3%
Betsiboka	5.0%
Melaky	4.8%
Atsimo Andrefana	12.6%
Androy	0.3%
Anosy	9.6%
Menabe	10.2%
Diana	20.3%
Sava	7.1%

Source: EPM 2010

NOTES

1. We use per capita consumption as the welfare indicator in this analysis.
2. While the dependent variable in this analysis is household per capita consumption expenditures, throughout the analysis the less cumbersome terms *consumption* or *household consumption* are used.
3. We acknowledge the assumption that the same-data generating process may be violated in future periods, but because we include variables such as experience of a climatic or health shock and key prices, which fluctuate over time, this assumption is not as strong as it may at first appear.
4. The advantage of the random selection of subsets of data in this algorithm is that it de-correlates the trees from one another and also reserves a subset of the data, not used to build a given tree, for unbiased testing of the accuracy of the prediction. This out-of-sample testing error is known as the out-of-bag error.
5. In particular, the mean-squared error (MSE) measure of each variable's importance in a regression forest is measured by randomly perturbing the variable of interest and recording the extent to which the out-of-bag error differs from that found with the unperturbed data (Hastie, Tibshirani, and Friedman 2009). The differences are averaged across all trees and then divided by the standard deviation of the differences to produce a normalized measure of the increase in MSE (%IncMSE), comparable across all variables.
6. Following guidance from Harvest Choice, TLU were calculated as follows: tlu = 0.7*ox + 0.7*cow + 0.1*sheep + 0.1*goat + 0.2*pig + 0.01*chicken + 0.01*turkey + 0.01*duck + 0.01*goose + .001*rabbit.
7. The *are* is a local unit of area measurement; 1 *are* equals 100 square meters or 0.01 hectares.
8. The 2010 EPM household survey has modules on both production and consumption of multiple commodities, including rice. Each module includes kilograms produced and consumed of paddy rice and dehulled rice. Net sellers/buyers of paddy rice and net sellers/buyers of dehulled rice are identified separately by calculating the marketable surplus (*marketable surplus = production – consumption*) of these two commodities for each household.
9. Availability of each of these commodities is a binary variable indicating whether the commodity price was reported for a given community in the community survey. Where no price was reported for a given commodity, this variable is zero for that commodity. Where any price was reported for a given commodity, this variable is one for that commodity.
10. Causal inference of the effect of expanded electricity would require additional empirical methods, which would require either more integrated or experimental data.
11. Unfortunately, it is not possible to observe in the data whether the dehulled rice category is composed of imported or white rice varieties or both (and in what proportion).

REFERENCES

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Deaton, A. S. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, MD: Johns Hopkins University Press.

Deaton, A. S., and S. Zaidi. 1999. "Guidelines for Constructing Consumption = Aggregates for Welfare Analysis." Mimeo. Princeton, NJ: Princeton University.

Hastie, T., R. J. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. New York: Springer.

Kleinberg, J., S. Mullainathan, and Z. Obermeyer. 2015. "Prediction Policy Problems." *American Economic Review: Papers and Proceedings*, 105(5): 491–95.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by Random Forest." *R News* 2:18–22.

Therneau, T., B. Atkinson, and B. Ripley. 2015. *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-10. Phoenix, AZ: Mayo Foundation. http://CRAN.R-project.org/package=rpart.

CHAPTER 4

Labor Demand Estimation in Rural Madagascar: Shadow Wages and Allocative Inefficiency

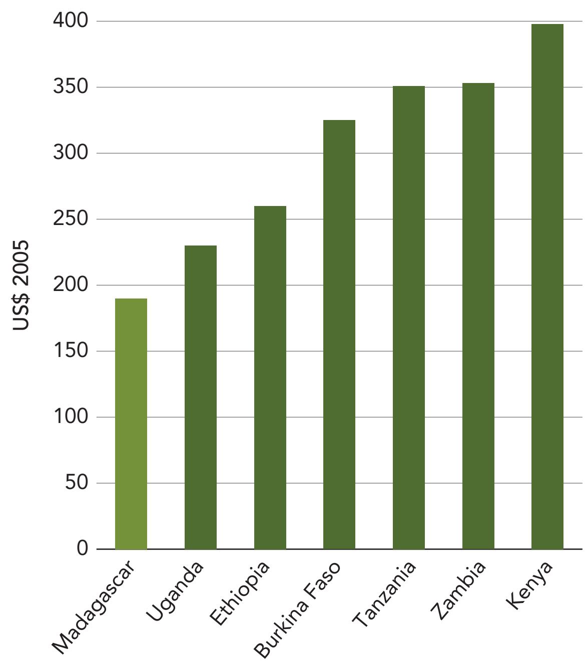
Margaret Jodlowski*
June 2016
*Cornell University
This research was conducted as part of a World Bank-supported analytical program under the task leadership of Theresa Osborne, who provided comments and guidance on earlier drafts, as well as the supervision of Christopher B. Barrett of Cornell University.

Introduction

Understanding the factors and circumstances that influence rural labor demand in Madagascar is of central importance for informing pro-poor growth policies. As with many poor countries, most of Madagascar's rural labor force is concentrated in agriculture and lives in poverty. Agricultural productivity is among the lowest in the world (see figure 4.1). Agricultural wage laborers, who tend to be among the poorest of the poor, are typically underemployed and paid very little. Many poor people work in both farm and rural nonfarm enterprises (NFEs), which have been shown to reduce rural poverty in poor countries (Barrett, Reardon, and Webb 2001), and yet these workers remain poor. A more expansive and efficient labor market is of key importance, not only for job creation and wages, but also for the productivity of farm and nonfarm enterprises. Thus, understanding the determinants of labor demand, including of any inefficiencies in these markets, is an important priority for designing pro-poor policies.

The major employer of labor in rural Madagascar is the household—households employ labor (household members and hired laborers) on farm plots and in household-operated NFEs, and they are also the suppliers of labor. Agricultural workers typically work on household farms,

FIGURE 4.1: Agriculture Value Added per Worker (Average 2011–14)



Source: World Development Indicators (WDI).

and the vast majority of those who earn income off-farm do so in the informal economy—working in an NFE, often receiving in-kind compensation, or being self-employed. More than 85 percent of Malagasy workers are employed in nonwage activities, and in 2005 only 11 percent of rural adults were employed as a nonfamily worker in an NFE (Stifel, Rakotomanana, and Celada 2007). Between 2001 and 2010, the percentage of households operating an NFE increased from 26.3 percent to 43.9 percent, while the percentage of these households that employed hired labor in their NFE declined from 30.8 percent in 2001 to 14.3 percent in 2005 and then stayed relatively constant between at 16 percent in 2010. Thus, a movement into NFEs was not accompanied by a greater willingness to hire nonhousehold labor. If this trend occurred despite the higher profit potential from hiring such workers, it would suggest a friction on the demand side of these labor markets, which reduces both labor and enterprise incomes in rural areas. Thus, understanding these and other outcomes requires an understanding of the factors influencing a household's demand for labor.

Constraints to raising labor demand can arise through the effects on the profitability (marginal revenue product, MRP) of labor, or through frictions in the labor market, and thereby the level of employment relative to the efficient (profit-maximizing) level. Further, there is strong evidence from across the continent that agricultural factor markets, especially the land and labor markets, do not function competitively and are subject to market failure. These failures are of potentially diverse origin, and include poor infrastructure and labor supervision problems. Barrett and Dillon (2016) reject the hypothesis of a well-functioning, complete, and competitive labor market in five Sub-Saharan African countries.[1]

Despite the widespread presumption that labor markets in poor rural economies are inefficient, there is relatively little research on the determinants of labor demand in such settings (see Hammermesh 1996). Jacoby (1993) was one of the first papers to structurally estimate shadow wages as the marginal revenue product of labor, in the presence of an informal (or nonexistent) wage economy where shadow wages are determined within the household. He observed that traditional methods for analyzing the labor supply decisions of households, and other household-based time allocation models, are inappropriate for contexts where self-employment is ubiquitous and wage rates are not observable. The assumption that the observed wage rate for laborers is equivalent for those who are self-employed is similarly invalid, as those who work to earn a wage and those who work on the household farm are likely to differ in both observable and unobservable ways. Thus, Jacoby developed a method to estimate structural time allocation models for households in the absence of observed market wages. Barrett, Sherlund, and Adesina (2008) generalized Jacoby's approach to accommodate risk, search, and transactions costs, as well as occupational and location preferences.

This paper attempts to build an empirical understanding of the functioning of Madagascar's rural labor markets, while also deriving insights into the factors affecting the revenues of rural households. In particular, following Randrianarisoa, Barrett, and Stifel (2009), we estimate the proximate drivers of demand for labor by rural households over the decade 2001 to 2010 using the Enquête Périodique auprès des Ménages (EPM) for the years 2001, 2005, and 2010. Because farm and nonfarm labor demand may differ, we analyze each sector separately, using only 2001 for the on-farm sector due to data limitations in the subsequent surveys. We adapt the methods developed to study labor supply by Jacoby (1993) and Barrett, Sherlund, and Adesina (2008) to the problem of labor demand, using their approach to address the issue of unobserved wages. We also relax the assumption that the wage is equal to the marginal revenue product of labor. We examine the shadow wage—the wage firms (households in our case) would be willing to pay labor.

Shadow wages are composed of two elements, the marginal revenue product of labor and an allocative inefficiency factor (AIF) that captures the effects of the nonwage costs (or benefits) which firms see when employing workers. If the nonwage costs exceed the benefits, this adjustment pushes down the wage the employer is willing to pay. If there are nonwage benefits to employing workers, such as retaining high-quality workers, future training benefits, or employment as a means of sharing resources with workers, the AIF will result in a willingness to pay more than the marginal revenue product of labor. In our sample, we find that there are relatively few cases of the latter. Given these two elements, shadow wages can be impacted by a vast array of factors. Affecting the marginal revenue product of labor are technology; relative output and factor price movements (macroeconomic variables and levels of market integration); and the cost and availability of other inputs, including infrastructure services, all of which affect the marginal revenue product of labor. In addition there are the risks—costs of hiring, training, supervising, and letting go of workers—that affect the AIF. These can be affected by institutional arrangements and household characteristics affecting the ability to reduce these costs. We estimate the observable determinants of job creation (the extensive margin of labor demand growth), as well as the increase in hours worked for the same number of jobs (the intensive margin), in addition to the AIF in rural labor markets. In addition, we estimate the responsiveness (or elasticity) of households' total demand for labor, both paid and unpaid, with respect to shifts in the supply of labor (or other nondemand-side drivers of wages), as well as the efficiency of these labor markets.

Based on the estimated divergence between the marginal revenue product and wages paid for households paying wages, we find evidence of significant allocative inefficiency in rural labor markets. A wedge equal to approximately half the marginal revenue product in effect halves the households' willingness to pay for labor. The wedge also appears to be higher in the NFE sector than on-farm: at the most extreme, wages for NFE workers in 2001 are only about 10 percent of the marginal revenue product of labor, while standard economic theory would equate them. The finding of such a large divergence between the observed wage and the marginal revenue product of labor is not necessarily evidence of miscalculation on the part of households. This wedge is related to nonwage costs and risks of hiring workers, but it is an important factor affecting both the potential to generate labor income and the profitability of household farm and nonfarm enterprises. Further, the allocative inefficiency we estimate is almost always negative: a negative value for the AIF indicates that labor is underdemanded by these household enterprises. This implies that there are barriers, only some which we can observe, to labor demand. We estimate the factors that are related to labor being over or under demanded and find that household enterprises for which the household head is well educated significantly over demand labor for both farms and NFEs in 2001. Also, the value of equipment was significantly related to an increased likelihood of labor being over demanded in 2001, but significantly related to an increased likelihood of labor being under demanded in 2010: capital investments into these small enterprises may be outpacing labor demand.

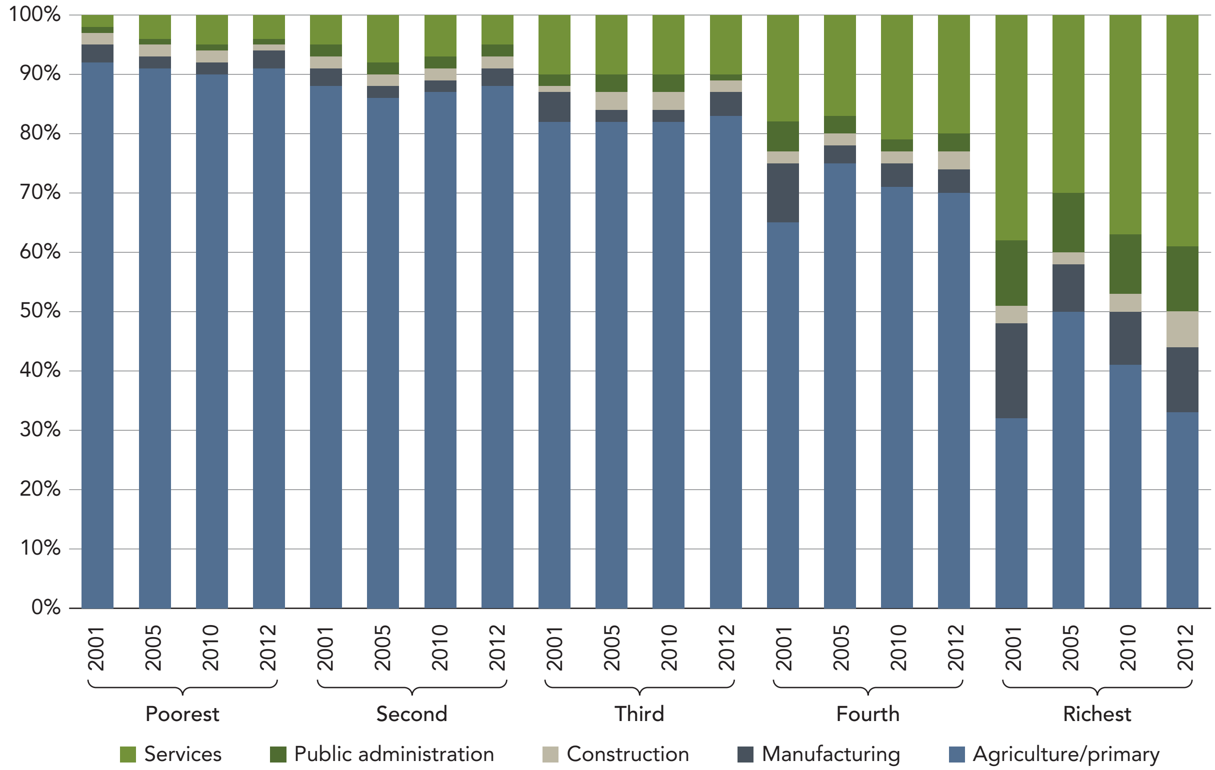
The availability and use of other inputs affects labor demand as well. We find that on-farm labor demand in 2001 is positively related to the land area cultivated and livestock holdings of the household, as might be expected.[2] For NFEs, a variety of factors significantly affect labor demand, with different ones emerging as important over time. First, the number of working-aged men and women in the household increased demand in 2001 and 2010, a likely result of the lower labor market frictions involved in employing family labor. In 2001, having more education increased labor demand, but in 2010 it reduced it. Our results highlight the importance of physical infrastructure for increasing the NFE revenue. In 2001 and 2005, higher transport costs are associated with lower levels of NFE revenue. In 2010, the availability of electricity and irrigation networks had a positive and significant relationship with revenue for NFEs, but they did not affect labor demand in any year studied, suggesting that labor market frictions are not helped by these services. Own investment in these small enterprises, measured by the value of equipment, also rises throughout the decade.

We also find that the demand for farm labor is wage elastic, while NFE labor demand is inelastic and becomes more inelastic over time. Based on the (Hicks-Marshall) theory of derived demand, elastic labor demand indicates that units of labor are easy to adjust with circumstances and workers are easily substituted, perhaps because the tasks performed by farm labor are neither highly specialized nor complex. For NFEs, however, this may not be the case. Rather than identify, hire, train, and supervise paid workers, NFEs prefer to utilize less labor and accept lower profits, given these nonwage costs. As a result, they tend to generate employment only of the household members and are not currently promising candidates for providing wage labor in rural areas.

Background and Data

As a country with a poverty rate of 77.8 percent, the first decade of the millennium was not kind to Madagascar.[3] The country experienced two political crises during the period covered by this study, first in 2002 and again in 2009, as well as a fiscal crisis in 2007. Labor markets were at least somewhat flexible in absorbing workers into different sectors as macro-conditions and policy responses changed.

FIGURE 4.2: Sector of Primary Employment

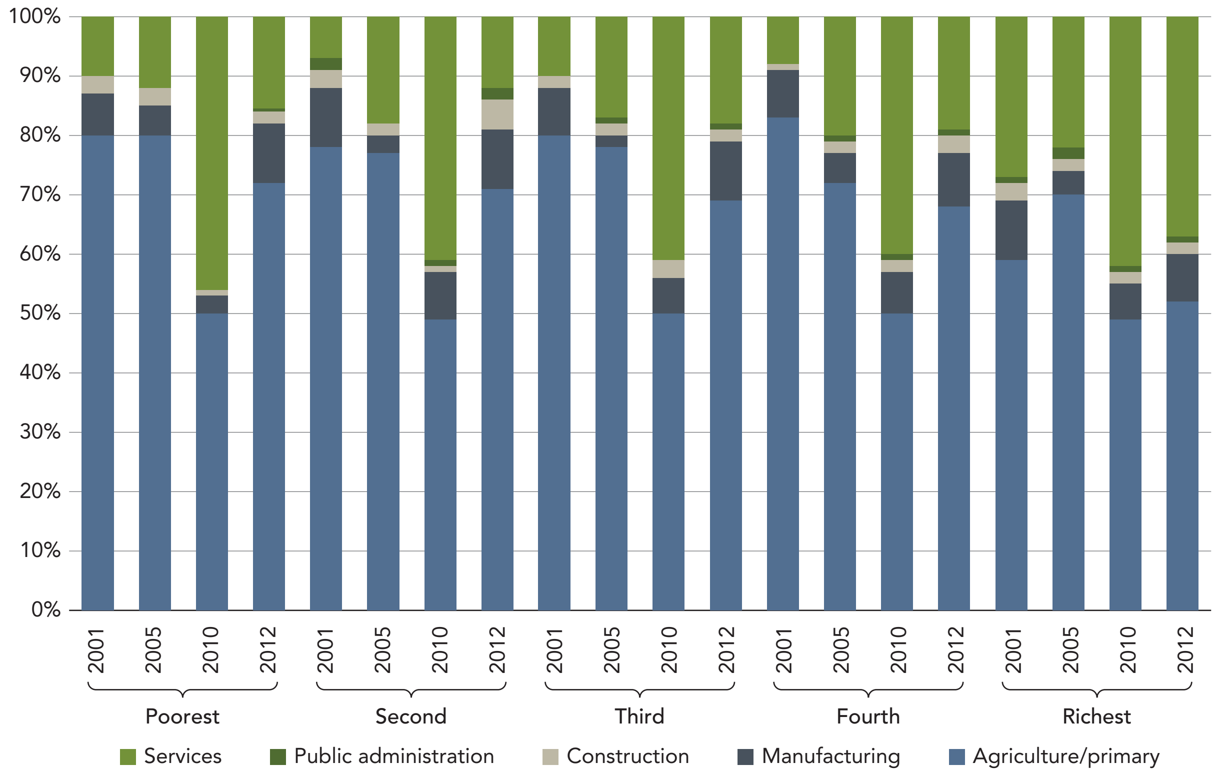


The decade saw people shift employment out of manufacturing and services and into agriculture (in 2005), and there was some evidence of urban-to-rural migration. Employment in agriculture by the country's relatively non-poor increased 24.8 percent from 2001 to 2005, while employment in manufacturing and services declined by 8.9 percent and 14.9 percent, respectively for the same group (Stifel, Rakotomanana, and Celada 2007). For those in the richest two income quintiles, primary employment in agriculture peaked in 2005, with the richest quintile making the largest jump, increasing from 32 percent employed in agriculture in 2001 to nearly 50 percent in 2005. However, by 2012, employment in agriculture had fallen relative to 2005, nonetheless remaining higher for the middle three quintiles than its 2001 level (figure 4.2). Across all income quintiles, the number of people whose secondary employment is in services increased markedly in 2010, with the largest increase coming from the poorest (from about 10 percent to nearly 50 percent), but this secondary service sector growth was not sustained. By 2012, secondary employment in services had fallen back to 2005 levels for all but the richest quintile, which saw a more modest decline (see figure 4.3). These shifts suggest that the profitability and opportunities in different sectors were subject to a variety of economic and other shocks (see, for example, Belghith, Randriankolona, and Osborne 2016; and Thiebaud, Osborne, and Belghith 2016).

Data

The data used in this paper are from the three most recent waves of the EPM (2001, 2005, and 2010). While the core modules for NFEs stayed the same throughout the three waves, funding and time constraints prevented the fielding of a detailed agricultural module in the years after 2001. Consequently, this analysis uses the estimates from the NFE data to describe the dynamics of labor-demand elasticities over time, while providing a static estimation of the shadow wage elasticity of demand for agricultural labor in 2001. The EPM surveys followed a two-stage stratification

FIGURE 4.3: Sector of Secondary Employment



procedure, with the first stratification being at the *faritany* or province level, while the second divided areas within these provinces into urban areas and rural areas. For further distinction, major cities (*grand centres urbain*) were differentiated from smaller urban centers (*centres urbain secondaire*). The agriculture section of the survey from the 2001 survey asks respondents about each plot they cultivate separately, including information about family labor, wage labor, and animal labor for each crop they report having cultivated. In each survey, households reported the number of workers (both family members and nonfamily members), equipment values, and various expenses for each NFE operated by a member of the household. These data, supplemented by household demographic information, form the basis of our analysis. In addition to the household survey data from the EPM, there are community-level data for each round, providing information on many community-level factors. We chose those that could influence the demand for labor as well as the cost of searching for and hiring labor.

As a measure of remoteness, we utilize the transportation cost of shipping a 50 kilogram bag of rice to the nearest main urban center during the rainy season. This proxies for the expense both of sourcing inputs to production, as well as the cost of marketing any agricultural surplus, or, in some cases, the final product of the NFE. Any substitutes for labor in the production function, especially modern ones like machinery or chemical pesticides or fertilizers, will be imported via the nearest urban center. The other community-level controls include an indicator of whether the community is considered part of a *zone rouge*, indicating a high level of crime and insecurity; whether the community has national television and radio coverage; and an indicator of whether the community has access to rural financial services. Especially in this tumultuous decade, physical security was likely to be a major driver of labor supply movements, with people avoiding more violent areas. Access to finance and broadcast media indicates the extent to which a commune is able to invest in their enterprises and keep abreast of market conditions.

Madagascar's Agricultural Sector

Madagascar's agricultural sector is characterized by the dominant production of rice, the country's staple grain, and of several nonrice food crops, grown both for home consumption and market sale. In addition, export crop production is concentrated in coastal areas and includes commodities such as vanilla, coffee, cocoa, and spices. Rice production uses more labor than any other kind of production, with an average of 54.08 person-days per hectare; the next most labor intensive crop types are export crops, which use 33.15 person-days on average. Because of its labor intensity, 32 percent of plots growing rice used hired labor, compared to only 5 percent of plots growing export crops. However, there is no significant difference between the wages paid to hired workers based on the crop type. Even for the most labor-intensive production, most of the labor comes from the family, from 84.6 percent for rice to 95.5 percent for export crops. Across all growing types, plot ownership rates are around 90 percent. However, some factors do seem to be correlated with an increase in the amount of hired labor used, especially education. Farm operators who have not completed secondary school—those with no education or some primary education (78.8 percent of operators in the sample) or who completed primary but not farther (14.7 percent)—hire 4.35 person-days of nonfamily labor; whereas those who have completed secondary school (5 percent), or have post-secondary education (1.5 percent), hire almost exactly double that amount, 8.68 person-days, on average, a statistically significant difference.

Summary statistics comparing plots operated by female- and male-headed households can be found below in annex 4B. There is no significant difference in the amount of hired labor employed by female-headed households, which constitutes 16.1 percent of the sample. There is also no significant difference in plot ownership rates, or even in the average area cultivated by households headed by either gender. However, male-headed households use significantly more rented tractor-hours and animal traction, both household-owned and rented. Additionally, male-headed households purchase significantly more pesticide and apply significantly more organic fertilizer, which could be a direct result of the higher amount of animal traction. Farm revenue function estimates (see table 4A.4) show that this seems to be driven by complementarity of the animal traction and NPK fertilizer inputs, the interaction of which is positive and significant in the farm gross revenue estimation. It seems, therefore, that while female-headed households face differential levels of access to agricultural inputs that are complements or substitutes to labor, they do not employ significantly different quantities of labor, either from their own family or hired.[4]

Agricultural producers choose both the extensive and intensive margin of input use from a diverse set of input choices. For tractability, we chose six inputs out of this set that saw the most widespread use across the sample; these are summarized in table 4.1, disaggregated by the type of crop.[5] Rice, the most commonly grown crop, uses significantly more labor (both family and hired) as well as more tractor-hours than either non-rice food crops or export crops. NPK fertilizer use is not common and does not differ across crop types. Labor, especially family labor, is the most commonly used input, regardless of the crop type, apart from land.

NFEs in Rural Madagascar

The NFEs described in this survey are small-scale businesses, largely operating in the informal sector and characterized by their small size. Each employs only 1.5 people on average (whether hired worker or household member) and earn an annual revenue of MGA 8.6 million, or US$1,448. There is, understandably, a good deal of heterogeneity in the operation and structure of these firms over the four sectors (agriculture, manufacturing, services, and trade). Over 500 unique business types are recorded over the three EPM survey rounds, with weavers and seamstresses being the most common. Grocers and other vendors are the next most frequent. Agricultural enterprises, which grew from 4 percent of enterprises in 2001 to nearly 25 percent in 2010, are the most informal (only 6 percent registered with the government, compared to 25.5 percent of trade enterprises and 27.8 percent of those in services) and operate fewer months out of the year (8.5 months) than the other types of enterprises (10.75 months for services, 10.64 for trade, and 9.31 for manufacturing). Thus, these are especially small and informal operations in an economy dominated by informality.

Over the decade, the percentage of surveyed households operating an NFE increased, while the percentage of

TABLE 4.1: Agricultural Inputs by Crop Type (Means, 2001)

	Rice	Nonrice food crops	Export crops
Area (in *ares*)	68.89 (114.6)	46.30*** (101.8)	66.15 (89.87)
Family labor (days)	51.56 (79.56)	25.76*** (49.57)	43.94** (70.40)
Hired labor (days)	15.04 (46.67)	7.895*** (40.77)	2.986*** (7.745)
Animal, own (hours)	39.24 (403.6)	17.24*** (216.0)	0.0485*** (0.547)
Tractor, own (hours)	24.75 (436.4)	11.46 (279.3)	0.0544*** (0.515)
NPK fertilizer (kg)	2.831 (46.29)	1.928 (66.28)	0*** (0)
TLU	2.439 (6.945)	3.599*** (10.80)	1.054*** (2.221)
Equipment value, log	3.362 (1.318)	3.272*** (1.347)	3.316 (1.170)
N	3,479	2,814	423

Note: Standard deviation in parentheses. TLU = tropical livestock unit. The *are* is a local unit of area measurement: 1 are = .01 hectare.
****Significantly different from rice at 1%.* **Significantly different from rice at 5%. *Significantly different from rice at 10% levels.

these households that employed hired labor in their NFE declined from 2001 to 2005, and then stayed relatively constant between 2005 and 2010. During the same time, households' investment in NFEs, measured by the value of their equipment, also increased. In 2001, of the 5,080 households surveyed, 1,334 operated a NFE (26.3 percent), and of those enterprises in operation, only 411 hired any nonfamily labor (30.8 percent of households operating a NFE). This percentage increased slightly to 30.3 percent of households operating a NFE in 2005, although fewer enterprises use hired labor in this year (14.3 percent of households operating a NFE). Finally, in 2010, 43.9 percent of households operated a NFE, although a similar percentage (16.0 percent) used hired labor. The drop in the percentage of households hiring labor could be a result of increasing household size, as shown in table 4.2. As households in rural and secondary urban areas grow, these family members can replace hired workers, and, given the increase in the number of NFEs in the sample, may start operating small enterprises of their own.

Table 4.3 shows that the composition and structure of NFEs change over time. Since NFEs were identified through household sampling, the analysis is not representative of larger firms. Only 10 percent of these NFEs in the sample, across all years, are registered with the government. Although the number of NFEs increases from year to year, the average number operated by a single household declines between 2001, 2005, and 2010, and of those, a decreasing percent respond yes to the question "Does the enterprise still have actual activity?" This could indicate that more households had started operating NFEs between 2005 and 2010, but that by the time they were surveyed in the last round, these operations had ceased their activities. Over the same period, the average amount of hourly wages paid to both family and hired workers increased from MGA 8.75 and MGA 52.51 in 2005 to MGA 10.15 and 59.54 in 2010, respectively. In 2001, NFE operators paid an average wage of MGA 30.00 between hired workers and family workers, whose wages were not reported separately in the survey. The value of equipment owned by NFEs increases

TABLE 4.2: Household Composition Changes

	2001	2005	2010
Men in household	1.078 (0.024)	1.107 (0.014)	1.174[a,b] (0.011)
Women in household	1.018 (0.022)	1.297[a] (0.012)	1.259[a,b] (0.009)

Note: Standard errors in parentheses.
[a]Significantly different from 2001.
[b]Significantly different from 2005.

TABLE 4.3: NFE Summary Statistics (by Year)

	2001 mean	2005 mean	2010 mean
Enterprise has had actual activity in the last year (1 = yes)*	0.979 (0.143)	0.961[a] (0.209)	0.953[a,b] (0.212)
Wage paid to household members (MGA/day)	—	8.751 (161.2)	10.15 (97.81)
Wage paid to hired workers (MGA/day)	—	52.51 (463.3)	59.54 (484.9)
Received financial aid	0.011 (0.002)	0.013 (0.002)	0.018[a,b] (0.002)
Number of household employees	1.231 (0.91)	1.486[a] (0.931)	1.634[a,b] (1.114)
Number of hired employees	0.356 (1.645)	0.345 (1.668)	0.451[a] (1.442)
Value of equipment (10,000 MGA)	111.8 (6965)	187.5[a] (2,905)	264.8[a] (2,945)
Years in operation	6.166 (9.495)	6.369 (7.468)	8.285[a,b] (8.832)
Number of enterprises operated by a household	1.321 (0.536)	1.211[a] (0.448)	1.118[a,b] (0.368)
Agriculture enterprise (1 = yes)	0.0523 (0.223)	0.042 (0.201)	0.28[a,b] (0.449)
Manufacturing enterprise (1 = yes)	0.154 (0.361)	0.024[a] (0.153)	0.0422[a] (0.201)
Trade enterprise (1 = yes)	0.174 (0.379)	0.0723[a] (0.259)	0.342[a,b] (0.474)
Services enterprise (1 = yes)	0.481 (0.500)	0.52[a] (0.500)	0.336[a,b] (0.472)
Monthly wages paid (MGA)	30.00 (189.75)	—	—
N	1,568	3,333	5,783

Note: Standard deviation in parentheses.
[a]Significantly different from 2001. [b]Significantly different from 2005.

markedly as well: from MGA 1.118 million in 2001 to MGA 1.875 million in 2005 and MGA 2.648 million in 2010.[6] Within a given year, there are more family workers on average in each NFE than hired workers, and they, unsurprisingly, receive less in wages than their nonfamily counterparts.

Table 4.4 shows the summary statistics for these community-level characteristics in 2001, as well as the summary statistics for some plot-level characteristics that impact production but are not inputs chosen by the operator. Most plots in the sample face some sort of disadvantage: erosion is the most common, with nearly three-fourths of plots considered eroded. Communities also face a lack of access to finance: less than 10 percent of communities have access to some sort of rural financial institution. Access to broadcast media is more common, but by no means universal: more than half of farms in the sample are in communities with no access to national television or radio.

Traditional labor supply models make assumptions that are empirically intractable, especially in places with limited formal labor markets, such as rural Madagascar. As a result, Barrett, Sherlund, and Adesina (2008) developed an extension of Jacoby (1993) in light of the implausibility of one of the original model's assumptions. As with many other early labor supply models, Jacoby

TABLE 4.4: Plot and Community Characteristics for Farms (2001)

	Mean
Plot characteristics	
Hillside (1 = plot is on the hillside)	.217 (.412)
Hilltop (1 = plot is on the top of a hill)	.116 (.32)
Eroded (1 = plot is eroded)	.732 (.443)
Sandy (1 = plot soil is sandy)	.146 (.353)
Pest (1 = plot experienced a pest attack in the last year)	.372 (.483)
Weather (1 = plot experienced a weather shock in the last year)	.516 (.5)
Community characteristics	
Transport cost	10366 (14,015)
Zone rouge (1 = yes)	.154 (.361)
Access to broadcast media (1 = yes)	.431 (.495)
Access to finance (1 = yes)	.0704 (.256)
N	7,671

Note: Standard deviation in parentheses. Transport costs reflect cost of transporting 50 kilograms of rice to nearest urban center in the rainy season.

assumed the textbook equilibrium condition of $MRP_L = w$, that the market wage is equal to the marginal revenue product of labor. Empirically, there are numerous reasons why this condition will be violated, including risk, search, and enforcement costs, among many others; statistically, papers that use Jacoby's method to structurally estimate labor supply routinely reject the hypothesis of equality: for example, Jacoby (1993) in Peru; Barrett, Sherlund, and Adesina (2008) in Côte d'Ivoire; and Skoufias (1994) in India. The deviation between MRP_L and w is defined as naïve allocative inefficiency (AI). Here, naïve reflects that this inefficiency is relative to a naïve model where such a deviation does not exist. The existence of AI is not necessarily an indication of error on the part of hiring households. In light of this, Barrett, Sherlund, and Adesina (2008) propose a method that takes into account the nonobservability of wages and of allocative inefficiencies for most households.

There are complications to empirically analyzing determinants of rural labor demand. The first is related to the structure of the data sets themselves, and the others are related to the data that are not observable. First, as discussed, the EPM data set is repeated cross-sections, rather than panel data. This complicates matters significantly, as there are distinct benefits to panel data, specifically in the ability to control for unobserved heterogeneity.

Another factor complicating demand estimation in this context is that the majority of labor demand in rural Madagascar comes from family enterprises, in which workers are not compensated with a wage but rather via a share of the profits or other in-kind remuneration. As a result, researchers either must make very strong assumptions, such as that those working on a family farm would be paid the same as an observationally similar individual with a recorded wage rate. Alternatively, researchers must impute a wage rate for these individuals.

Wages, when observed and recorded, often represent only the recorded cost of hiring a worker, although there are many other costs associated with the hiring and maintenance of staff. The researcher does not observe searching, hiring, monitoring, or supervision costs, or the costs of firing workers, even when wages are recorded. Yet employers factor these costs into hiring and wage decisions. In this context, such costs are likely to exhibit a good deal of heterogeneity related to whether these costs are borne by an agricultural enterprise or a NFE, as well as related to location and attributes of the employer and enterprise. For example, Otsuka and Yamano (2006) point out, "The cost of monitoring the work efforts of [agricultural] laborers in ecologically diverse farm environments is exceedingly high." As a result, labor is often only demanded for tasks that do not require much skill or are easy to monitor, when, in the absence of these high monitoring costs, demanding additional labor for more specialized tasks could be revenue- and welfare-improving. In order to estimate labor demand parameters consistently, therefore, we must control for this systematic variation in the "true" shadow cost of employing labor.

Although in this context labor may move without restriction between on- and off-farm employment, the two sectors may differ appreciably in terms of whether they are affected by seasonality and weather shocks. As a result, there are likely to be structural differences between in terms of labor demand patterns. To address this issue, we estimate labor demand in the two sectors separately. Because of changes in the survey, we are only able to accomplish this for the 2001 round of the EPM, as data on agricultural inputs, including labor, were not collected in later survey rounds.

Empirical Strategy

The empirical strategy we implement is designed to parameterize the household-level conditional-factor demand functions for labor in farms and NFEs, while addressing the theoretical and data challenges described. Estimation is theoretically motivated by an enterprise-based household model, in which households choose consumption of home-produced and market goods; labor allocation among leisure, home production, and wage employment for each household member; and whether or not to hire nonfamily labor to supplement or replace family labor. The enterprise the household operates can be either a farm or a NFE. The household makes these decisions in order to maximize household utility, subject to a budget constraint.

A household's hiring decisions are nonseparable from their consumption and their labor market participation decisions. This nonseparability arises because family and nonfamily labor are not perfect substitutes, due to supervision and search costs of hired labor, risk premia, and liquidity constraints. Therefore, household demographics, in addition to standard firm characteristics that might affect these sources of friction, must be considered when estimating labor demand.

Estimation follows a four-step procedure, as outlined in Barrett, Sherlund, and Adesina (2008). The final step contains the primary model of interest: household-level demand for labor in rural Madagascar. Because labor employed is censored at zero, the main empirical model we estimate is a censored Tobit regression:

$$m_i^* = \beta_0 + \beta_1 w_i^* + \beta_2 A_i + \eta_i$$

such that:

$$m_i = \begin{cases} m_i^* & \text{if } m_i^* > 0 \\ 0 & \text{if } m_i^* \leq 0, \end{cases} \tag{1}$$

where m_i^* is the equilibrium amount of labor employed by household i, m_i is the observed level of labor employed, w_i^* is the shadow wage rate, which itself must be estimated, and A_i is a vector of household and enterprise characteristics, including the characteristics of the community in which the household resides. We assume that η_i, the error term, is normally distributed with mean zero, as required for Tobit maximum likelihood estimation.

The four steps required to estimate this final model are as follows:

1. First, one estimates the enterprise production function and recovers the implied marginal revenue product of labor.
2. For the subsample of enterprises that pay workers a wage, one estimates the enterprise-specific divergence between the observed wage rate paid to workers and the estimated MRP_L from step 1, as a function of enterprise, employer, and community attributes.
3. One calculates the shadow wage, w^*, for all enterprises by adjusting the estimated MRP_L for the estimated AIF from step 2.
4. One estimates the labor demand function, equation (1).

STEP 1: ESTIMATING MRP_L

First, we use the entire sample to estimate stochastic revenue functions for both farm and nonfarm production.[7] The dependent variable in the first step is the annual revenue (gross revenue, minus expenses and salaries paid, for each NFE operated by a household and gross agricultural revenue per plot for farm production)[8] in order to aggregate across the wide variety of products produced by households in the sample and ensure comparability between farms and NFEs. The regressors for farm production include the quantities of the main inputs to production: total labor, animal traction, tractor usage, NPK fertilizer, and land. To capture land quality, we also include controls for the plot-level characteristics listed in table 4.1. There are fewer observed inputs for the NFEs: those we use are total labor, the value of equipment, and the amount of financial aid received.[9] Controls for the number of years the NFE has been in operation, the sector it operates in, and whether it is reported to have had "actual" activity in the past 12 months are also included as controls.[10] This approach is not without its drawbacks, namely, the likely simultaneity of the input application rates, as revenue and input application rates may be affected by unobserved factors that violate the orthogonality condition for OLS estimates to be unbiased. Even when using household-level fixed effects, the data we have on plot-level characteristics do not capture the diversity of agronomic conditions that farm operators observe. Farmers use these conditions, such as soil composition and quality, drainage, slope, and location on the farm, when making decisions about application rates of other inputs, including labor. Also, inputs' application responds to unobserved shocks in the error term, η, (such as pests), also violating the orthogonality condition.

We estimate the production function using a generalized Leontief second-order flexible functional form which allows for the flexible identification of complementarities between inputs. While any second-order flexible function form provides an exact second-order approximation of the true, unknown function at the sample means, the generalized Leontief specification additionally allows for input values to be zero, as is often the case in this context, in contrast to the translog functional form (Chambers 1988). The generalized Leontief specification is as follows:

$$TR_k^{1/2} = \gamma_0 + \sum_{i=1}^{m} \gamma_{ik} x_{ik}^{1/2} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \gamma_{ij}\, x_{ik}^{1/2} x_{jk}^{1/2} + \Gamma Z_k + \epsilon_k, \quad (2)$$

where TR is total revenue for household-enterprise k; x_i is the quantity used for each of the m inputs; Z is a vector of plot, enterprise, and community characteristics that directly affect production; and ϵ is a mean zero independent and identically distributed (iid) error term.[11]

Denoting labor by subscript L, the estimated marginal revenue product of hired labor, $\widehat{MRP}_L$, for each household-enterprise can be estimated by taking the partial derivative of (2) with respect to labor, which we will call x_L, where $L \in [1, m]$, using the parameter estimates, $\hat{\gamma}_k$, of the γ_0 to γ_m terms above:

$$\widehat{MRP_L} = \widehat{\frac{\partial TR}{\partial x_L}} = \frac{TR}{x_L}\left[\hat{\gamma}_L + \sum_{j=1}^{m} x_L^{1/2} x_j^{1/2} \hat{\gamma}_{Lj}\right]. \quad (3)$$

STEP 2: ESTIMATING THE ALLOCATIVE INEFFICIENCY FACTOR

The naïve allocative AIF is estimated in the second step for each household-enterprise that hires wage labor in our sample.[12] Using the $\widehat{MRP}_L$ from step 1 and the observed wage w from the data set, AIF is defined within the subsample of household enterprises that paid workers by the following relationship:

$$AIF = \ln\left(\frac{w}{\widehat{MRP_L}}\right). \quad (4)$$

Since our hypothesis is that $w \le MRP_L$, we expect this expression will be negative. A negative AIF indicates that, relative to labor for market wages, on-farm (or on-enterprise) labor is under demanded; the opposite holds for a positive AIF value, which indicates labor is being over demanded. Of course, an AIF value of zero means that the wage equals the marginal revenue product of labor and so there is no naïve inefficiency. Because the $\widehat{MRP_L}$ may deviate systematically from w across the different enterprises we observe based on their characteristics, especially between those engaged in agriculture and those not engaged in agriculture, we attempt to identify the characteristics correlated with AIF by regressing it on H, a set of enterprise and operator characteristics. This regression takes the following form:

$$AIF = \alpha_0 + \alpha_1 H + \mu, \quad (5)$$

where μ is a mean zero iid error term. This set of characteristics H includes demographic variables, such as the number of working-age adults of each gender and the number of children in the household; characteristics of the household head, such as age, education, and migrant status; and community characteristics such as province-level dummies, access to transportation or financial services, and physical insecurity. These results provide correlates of AI, which can help us understand patterns of allocative inefficiency within subsets of the population. More crucially for this analysis, estimation of equation (5) yields predictions of $\widehat{AI}$ for the households that we do not observe hiring nonfamily labor.

STEP 3: IMPUTING SHADOW WAGES

The third step combines the estimated allocative inefficiency $\widehat{AI}$ from step 2 and the estimated $\widehat{MRP_L}$ from step 1 to impute shadow wages, $\hat{w}^*$, for all households by rearranging equation (4) to estimate:

$$\hat{w}^* = e^{\widehat{AT}} * M\hat{\hat{R}}P_L \tag{6}$$

This shadow wage constitutes a sufficient statistic to address the issue of nonseparability of household production and consumption decisions (Jacoby 1993).

STEP 4: ESTIMATING LABOR DEMAND

Taking the imputed shadow wage $\hat{w}^*$, we estimate labor demand as in equation (1). In step 4, the dependent variable is a household enterprise's latent demand for labor. We bootstrap the standard errors (with 500 replications) of the Tobit regressions in order to mitigate the problems produced by sequential, multistep regressions estimations such as this.[13]

The Tobit model gives only one point estimate for each coefficient, and so to isolate the change in the probability of using labor (the extensive margin) from the change in the amount of labor used (the intensive margin), we follow McDonald and Moffitt (1980), who were the first to propose this decomposition as an extension of the Tobit model.[14] We therefore report three separate marginal effects:

a. $\frac{\partial E(d_i)}{\partial x_i}$: change in the unconditional expectation of latent demand for labor

b. $\frac{\partial E(d_i | d_i > 0)}{\partial x_i}$: the change in the expected level of observed use conditional on the household actually using labor, or the intensive margin

c. $\frac{\partial P(d_i > 0)}{\partial x_i}$: the change in the probability of labor being used, or the extensive margin

Results

PRODUCTION FUNCTION AND MARGINAL REVENUE PRODUCT OF LABOR ESTIMATES

The full estimation results of the stochastic revenue function for agriculture (in 2001) appear in table 4A.4 and the results for NFEs' net revenue for 2001, 2005, and 2010 appear in table 4A.5 The main input variable of interest, total labor, is statistically significant and positive for farms and NFEs in 2001 and NFEs in 2010. However, additional labor was associated with a significant decrease in net revenue in NFEs in 2005, a year of serious disruption in urban labor markets. This may be the result of rigidities in the level of labor employed at a time of falling profitability. Other inputs that positively contributed to farm revenue include NPK fertilizer usage and household plot ownership versus other forms of land rights. This indicates that the use of modern inputs increases farm revenue, and that land tenure and security does as well. Households may be more likely to invest in the long-term productivity and health of their plot if they own it.

Our analysis of NFEs shows that increases in years in operation has a positive and significant relationship with NFE revenue in 2005 and 2010. A dummy variable that indicates that a household enterprise received financial aid is positively associated with revenue in 2010, indicating the importance of outside sources of capital for these small businesses. Such sources of financial aid include microfinance institutions, government grants, and, most frequently, financial support from friends and family, including remittances. There are, therefore, likely important network effects that help determine whether an NFE will have financial success in a given year. There are also important benefits, in terms of increased revenue, from better infrastructure: increasing transport costs have a negative relationship with NFE revenue (significant in all years except 2010). In 2010, access to electricity and irrigation systems (both measured at the community level, rather than the household level) improve NFE revenue as well. Finally, there is a significantly negative association with physical insecurity and revenue in 2005 for NFEs. Unsurprisingly perhaps, physical insecurity, theft, and violence are bad for business.

The elasticities and effects on the estimated marginal revenue products of labor are shown in table 4.5. The MRP elasticity for each input (total labor for both farms and NFEs, and also land and NPK for farms) was calculated by computing the elasticity for each household and then taking the average across these values, rather than computing the elasticity at the mean of each input variable. For farms, a 1 percent increase in labor used increases revenue by 0.30 percent; for NFEs in the same year a 1 percent increase in labor used increases net revenue by 0.18 percent. The elasticity of revenue with respect to labor use, however, declines over the next two sampled years: a 1 percent increase in labor increases

TABLE 4.5: Estimated Elasticities of Revenue and Marginal Revenue Product of Labor

Elasticities	NFE 2001	NFE 2005	NFE 2010	Farm 2001
Total labor	0.018 (0.006)	0.0081[a] (0.002)	0.009[a] (0.0015)	0.301 (0.801)
Land area	—	—	—	–0.072 (0.292)
NPK	—	—	—	0.001 (0.0039)
NFE Equipment	–0.078 (0.011)	–0.016 (0.0121)	–0.0284 (0.0041)	
Marginal revenue product				
Total labor (MGA)	60,647.2 (4001.2)	58,321.6 (1147.7)	27,231.7[a] (80876.5)	7,689.131 (1302.234)
Land (MGA)	—	—	—	–30,797.18 (432,491.5)
NPK (MGA)	—	—	—	828.488 (1987.6)
NFE equipment	–19,651.2 (7332.12)	–14,002.6[a] (9768.81)	–7454.57[a] (2618.7)	
AIF, means	–2.331 (.226)	–2.567[a] (.291)	–1.918[a] (.147)	–1.578 (.070)
AIF, medians	–1.502	–2.065[a]	–1.872[a]	–1.573

Note: Standard deviations in parentheses. NFE = nonfarm enterprise. NPK = fertilizer. AIF = allocative inefficiency factor.
[s] Significantly different from 2001 value.

revenue by 0.08 percent in 2005 and 0.09 percent in 2010. In the latter half of the decade, following the two political crises, the elasticity of revenue with respect to labor declines significantly. In 2001, the estimated marginal revenue product of labor for farms is significantly lower than for NFEs: the marginal revenue product of labor for farms in 2001 is about MGA 7,000: nearly ten times less than the marginal revenue product of labor for nonfarm enterprises in that year. But it is also significantly higher than the average wage paid to employees in the sample, which are lowest on average for NFE employees in 2005 and 2010 (about MGA 2,000 per day) and highest on average (MGA 7,000 per day) for NFE workers in 2001, with farmworkers in that year earning around MGA 5,000 per day. (See table 4.6 for the wage data).

TABLE 4.6: Estimated Shadow Wages and Observed Wages (MGA per Day)

Shadow wages	Farm, 2001 mean	NFE, 2001 mean	NFE, 2005 mean	NFE, 2010 mean
Nonhiring enterprise	7434.72 (2864.544)	7478.84 (25316.15)	3233.421[a] (5429.388)	3869.321[a] (9386.16)
Hiring enterprise	7512.422** (955.843)	10287.15*** (16112.61)	6945.05[a,**] (12015.4)	8008.15[a,**] (1817.816)
Observed wages	5652.614††† (927.656)	6867.779††† (32226.84)	2667.61[a,†††] (5445.34)	2049.57[a,†††] (14927.18)

Note: Standard deviation in parentheses. NFE = nonfarm enterprise.
***, **, *Significantly different from hiring at 1%, 5%, 10% levels, respectively.
†††,††,†Significantly different from MRP_L at 1%, 5%, and 10% levels, respectively.
[a]Significantly different from 2001.

ALLOCATIVE INEFFICIENCY FACTOR ESTIMATES

It is possible to test whether the textbook condition assumed by Jacoby (1993)—that the market wage equals the marginal revenue product of labor—is indeed violated in this case by comparing the observation-specific values for $M\hat{R}P_L$ estimated in step 1 with the observed wage for enterprises that hire workers (and, importantly, also record a wage). The test is a bivariate regression of log wages on log $M\hat{R}P_L$, with the null hypothesis that there is no allocative inefficiency (i.e., $\hat{\alpha} = 0$ and $\hat{\beta} = 1$):

$$\ln(w) = \alpha + \beta \ln(M\hat{\hat{R}}P_L) + \epsilon. \qquad (7)$$

We reject this null for all years. This reaffirms the finding of Barrett, Sherlund, and Adesina (2008), among others, and demonstrates the need to estimate the divergence between the observed wage and the $M\hat{R}P_L$ systematically. The marginal revenue product of labor must be adjusted for the unobserved costs of hiring workers to accurately estimate demand. To make this adjustment, we regress the AIF, as calculated from equation (4), on a set of household characteristics in order to recover correlates to use for more accurate estimation of the shadow wage. The full results of this estimation are found in table 4A.7. Across all four specifications, the AIF was negative 92.1 percent of the time, implying that the wages paid are consistently lower than the MRP of labor and that there are important nonwage costs of hiring labor for household enterprises. Further, a negative value of AIF indicates that labor is being underutilized relative to market wage work on these plots.

Table 4A.7 shows that the determinants of allocative inefficiency change from year to year for the NFEs. In 2001, for example, labor is more likely to be over demanded in NFE operations where the household head has at least some secondary school education, with no effect in later years. The likelihood of labor being over demanded relative to its MRP also increases with increased equipment in 2001 as well. In 2010, however, increased equipment value becomes significantly correlated with the likelihood that labor is being under demanded. NFE operators may not be able to hire at a rate that keeps pace with their increasing capital investments. In 2010, the year in which secondary employment in services increased dramatically across the income distribution, enterprises in the service industry were significantly more likely to have over demanded labor, as were manufacturing enterprises. Trade enterprises were more likely to have underutilized labor in this year. In 2005, the opposite is true: labor is underdemanded in service sector NFEs. As for farms in 2001, labor is under demanded on plots growing rice and non-rice food crops. Increased total household landholding and, once again, education of the household head are associated with an increased likelihood of over demanding labor. Households with well-educated heads may over demand labor because their educational advancement allows them, or perhaps even obligates them, to serve as an employment safety net for their families and communities. It is important to remember, however, that these results capture associations rather than causal relationships, and that the existence of allocative inefficiency, especially in such a context where the labor market is so informal, does not necessarily reflect operator error or misallocation. Instead, it reflects market-wide frictions that affect all households, with certain household characteristics being correlated with the increased likelihood of labor being over- or undersupplied.

SHADOW WAGE AND DEMAND FOR NONFAMILY LABOR ESTIMATES

Following equation (6), we estimate shadow wages—the willingness to pay a wage for hiring labor from the firm's perspective—for the whole sample, both hiring and nonhiring households, using the estimated $M\hat{R}P_L$ from step 1 and the estimated $\widehat{AIF}$ from step 2. Calculated shadow wages, for both hiring and nonhiring farm plots, are found in table 4.6, in MGA per day. For all four specifications, the shadow wage for the nonhiring enterprises is significantly lower than for the hiring enterprises, as expected. Among the hiring enterprises, the average annual mean of the shadow wages ranges from MGA 3,233 for an NFE in 2005 to MGA 7,479 for an NFE in 2001. The shadow wage falls significantly for NFEs from its 2001 high in both 2005 and 2010. This could reflect the existence of excess labor supply in rural areas, prompted by increased urban-to-rural migration. This is supported in part by data which show the number of working-age adults in the household increasing throughout the decade (see table 4.2).

Based on these reported wages and the national poverty lines for each year, wages in 2005 were actually highest, as measured by the number of days a person would have to work to meet the national poverty line. In 2001, the average NFE worker would have to work 144 days; in

2005, that number decreases to 114.5 days. It peaks in 2010, when workers would have to work 228.7 days to meet the national poverty line for that year. The average NFE hires less than half of a worker and operates nine months out of the year, in a given year. Labor usage by household NFEs or farms is therefore not high enough to be a viable route out of poverty. Labor is underutilized from a firm perspective as well, based on estimates of the allocative inefficiency. Based on the changes in the marginal revenue product of labor, these additional workers benefit the firms that employ them, though at a falling rate over the decade: the MRP_L for NFEs is less than half its 2001 value in 2010. It could be that these household firms first absorb the labor offered by family members, as the number of family employees increases over time. The number of hired, nonfamily employees is not significantly different between 2001 and 2005, but the 2010 level is significantly greater. Together, these trends indicate that labor utilization increases as excess supply pushes wages down. For the sake of comparison, Barrett, Sherlund, and Adesina (2008) found observed wages for workers on Ivorian rice plots to be 56.7 percent of the marginal revenue product of labor there; in our case we find observed farmworker wages to be 73.4 percent of the MRP_L and NFE worker wages to be only 11.4 percent, in 2001. Thus, we see greater evidence of market failures in the NFE sector, as wages there are much farther from estimates of the marginal revenue product of labor. This suggests greater difficulty in finding appropriate labor for NFEs, compared to farms, where the work is likely to be less specialized.

Table 4A.9 presents the estimated marginal effects from the Tobit regression of demand for labor and the bootstrapped standard errors for the farms; these results for NFEs can be found in table 4A.10. Column 1 presents the estimate for the unconditional demand for labor, while columns 2 and 3 present the intensive marginal and extensive marginal labor demand effect, respectively, following a McDonald and Moffitt (1980) decomposition. We find that, controlling for community and household characteristics, the quantity of labor demanded falls as shadow wages rise, as expected. We find that labor demand for farm work is elastic, but that the opposite holds for nonfarm work: NFE labor demand is inelastic, and becomes more inelastic over time. The intensive margin dominates the extensive one for all cases except NFEs in 2001, implying that the elasticity of labor demand is being driven by the intensive margin. If, therefore, shadow wages were to suddenly to increase, and employing workers were to become more costly, firms would respond by reducing the intensity at which their current employees work, rather than reducing the number of employees they have. In 2001, on the other hand, the extensive margin dominates, meaning that changes in demand shifters were more likely to change the probability that additional labor was used, rather than increasing the intensity of work for currently employed labor. Because the MRP_L and the allocative efficiency both contribute to determining the shadow wage and factors that influence one over the other, it is not possible to estimate their effects on the labor demand separately from each other.

For farms, the most economically significant demand shifter is the education level of the household head (table 4.7). Specifically, a household head having some high school increases the amount of labor demanded by 34 percent. The effect is not seen for higher levels of education, but that could be a result of very few household heads having completed high school or postsecondary education. Labor demand also increases in the number of TLUs, which indicates that either farm animals, such as larger draft animals, are complementary to labor, rather than a substitute for it, or that larger herds of smaller animals or chickens require more labor to care for them. Similarly, increased landholdings also increase the amount of labor demanded. Of the community characteristics, being in a *zone rouge* area increases labor demand. This could indicate that although the cost of traveling to a job is higher, those households in areas with greater levels of physical insecurity may hire more labor to serve as guards for crops and livestock.

The results for NFEs show important changes in terms of what affects labor demand over time. In both 2001 and 2010, the number of working-age men and working-age women in a household is related to a significant increase in the amount of labor demanded, with an insignificant coefficient in 2005. This was the year in which household size was the largest and allocative inefficiency was the most negative on average, which indicates labor is being undersupplied. The number of family members employed in this year is significantly greater than in 2001: households are growing and these members are being put to work, but labor remains undersupplied, meaning that labor market frictions remain costly. Because most of the increase in household size comes from additional adult women and children, however, the new household

TABLE 4.7: Farms: Estimated Demand for Labor (Select Results)

	Farm 2001		
	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)
Shadow wage (ln MGA)	–1.599***	–1.324***	–0.1024***
	(0.107)	(0.085)	(0.01463)
Some high school	0.340*	0.3230*	0.0182**
	(0.198)	(0.1923)	(0.0078)
Zone rouge	0.657***	0.6276***	0.0325***
	(0.121)	(0.1198)	(0.0068)
TLU	0.0196***	0.1358***	0.0013***
	(0.00709)	(0.0295)	(0.0005)
Total land (in Ares)	0.162***	–0.444***	0.0105***
	(0.0339)	(0.0785)	(0.0023)
Rice	–0.531***	–0.808***	–0.0434***
	(0.0933)	(0.1301)	(0.01061)
Export crop	–1.213***	0.2814	–0.0374
	(0.175)	(0.24)	(0.03169)
District (fixed) effects	Yes	Yes	Yes
Constant	20.69***		
	(1.484)		
σ	1.032***		
	(0.0334)		
n	2,390	2,390	2,390

Note: Standard errors in parentheses. *Unconditional* refers to the unconditional expectation of the observed dependent variable. TLU = tropical livestock unit. ln=natural logarithm.
***p < 0.01, **p < 0.05, *p < 0.1

members could be ill suited for the work in the household's NFEs. The additional burden of supplying for a growing household might drive many NFEs out of business, reducing their quantity demanded for labor in this year. Additionally, the number of children becomes significantly related to quantity of labor demanded in 2010, indicating that household size and composition becomes increasingly important throughout the decade, perhaps as rural households grow as a result of increased urban-to-rural migration. Also different between 2005 and the other two years is the relationship between quantity of labor demanded and the gender of the household head. In 2001, male household heads demanded less labor, while in 2005 they demanded significantly more. Here, the household having a male head is associated with a 27 percent increase in the quantity of labor demanded.

Interestingly, the role played by education is not consistent. In 2001, the impact of education was positive. Completing primary school was associated with a 30.7 percent increase in the quantity of labor demanded, and having completed secondary school with a 24.8 percent increase, making educational attainment the most economically significant demand shifter in this year. However, while the relationship was still positive in 2005, it was not significant, and the relationship actually became negative in 2010, although the magnitude is small compared to 2001 and significant only for some high school and completed high school. This could indicate that while educational attainment is important, it is not enough to increase the quantity of labor demanded in the adverse economic conditions created by the 2009 political crisis. In light of such instability, few household characteristics are important.

Another potential impact of the 2009 crisis can be seen in the coefficients for transport cost. For 2010 it is actually positive and significant. Thus, being in an area that is more remote is actually associated with an increase in the quantity of labor demanded. These remote areas also were those least likely to feel the direct impact of the crisis, as they are farther from the country's urban areas, where political violence was concentrated. The *zone rouge* indicator is insignificant in 2005 and 2010, showing that physical violence and insecurity may have started to be associated with lower levels of labor demanded. As with farms, physical insecurity, measured by this indicator, is significant and positive in 2001. Finally, the relationship of equipment value to labor demanded changed markedly over time as well. While increased value of equipment was associated with increased quantities of labor demanded in 2001 and 2005, indicating that equipment is complementary to the activities of labor or that the equipment requires workers to maintain and service it, the relationship becomes significantly negative in 2010, which is also when the stock of equipment owned is also the highest. This suggests that by the end of the decade, capital has started to become a substitute for labor in these NFEs. The relationship between sector-level indicators and labor demand is not consistent year to year, indicating that no one sector experienced consistent growth or decline of labor demanded. For example, in 2001, manufacturing enterprises demanded significantly more labor, while in 2010 they demanded significantly less. In fact, in 2010, all sectors (agriculture, trade, and manufacturing) demanded significantly less labor relative to services, which is consistent with the service sector growth shown in figures 4.2 and 4.3 for the latter half of the decade.

Discussion and Conclusion

This paper estimated which factors increased (or decreased) rural labor demand for both on-farm and off-farm work by households in Madagascar, during a politically and economically turbulent decade. This estimation took into account the inherent nonobservability of wages for most informal sector or self-employed workers, as well as the unobserved nonwage costs of hiring and employing labor, which we found to be about half (51.8 percent) of the total cost of demanding labor in this context. These costs, which represent a variety of unobserved labor market frictions, are significant and make it more difficult for households to employ workers.

Indeed, the results of our analysis show that labor is undersupplied on more than 90 percent of household plots or in household NFEs. For all of these households, the observed wage is consistently and significantly lower than the marginal revenue product of labor. In order for labor to be efficiently and effectively allocated for both profit maximization and poverty reduction, the labor market frictions that inflate the cost of hiring and maintaining workers for employer households should be reduced. Because these frictions are unobserved almost by nature, specific prescriptions for their removal are difficult to make. Nonetheless, lessons can be applied from places where we observe well-functioning labor markets and work that specifically tests for the presence of labor market failures. Complete and competitive labor markets rely on infrastructure that facilitates easy job searches, have processes in place to write and enforce contracts, and have unrestricted worker mobility. Worker mobility is, at least in part, a function of a stable and peaceful society, where travel and relocation are not limited by fears of violence or unrest along the route. Especially in the specific context of Madagascar, where cultural ties to specific places and plots of land are strong, a lack of worker mobility may be especially widespread (Stifel, Fafchamps, and Minten 2011). Systems that help mitigate output risk, especially in agriculture, would also help labor markets function: concerns about crop failures or price shocks keep farmers from using the optimal level of all inputs, including labor.

Nonetheless, despite the high cost of employing workers for households, we do find some evidence of which factors are related to an increased quantity demanded of labor. We find that, especially for farms, educational attainment of household heads can stimulate rural labor demand, a positive externality of educational attainment. Indeed, education has benefits on the supply side of the labor market as well, as it allows workers to access better opportunities and attenuate their exposure to labor market risks (IFAD 2011). The positive relationship between land holdings and livestock holdings on quantity of labor demanded indicates that asset accumulation, especially productive assets like these, are beneficial for the local labor market, as these results suggest the two are complements rather than substitutes: having more land or more livestock elicits a positive labor demand response from farm households.[15]

For the informal enterprise sector, capital investments (measured by the value of equipment) have a changing relationship with labor demand over time. While the two demonstrate complementarity in the first two years studied (2001 and 2005), with the value of equipment associated with increased levels of labor demand, this relationship reverses in 2010. Interestingly, more remote areas are associated with higher labor demand in 2010 as well, and there is no longer a detectable positive externality from physical insecurity. One resource not covered in the survey, perhaps because it is unlikely to exist, is access to business development services. Such services can better equip microentrepreneurs to build their capabilities in all areas of business management, including labor relations. Technical and vocational skills development is another overlooked but potentially important way of improving engagement between the demands of the nonfarm labor market and the skills of the population in rural areas (IFAD 2011).

Finally, our results imply that exogenous wage growth (or wage growth due to large urban-center-based labor demand) would have markedly different effects on farm versus nonfarm employment. Based on the estimated elasticities, wage increases would have only a small effect in reducing the quantity of labor demanded in the more rural NFE sector, but the opposite is true in the farm sector. This higher wage responsiveness could reflect the thin margins typically observed in agriculture, whereby small increases in costs would result in a greater adjustment of the relevant input. However, in both the farm and nonfarm sectors, significant frictions exist. Wages paid in this environment are routinely significantly less than the MRP of labor. By identifying and reducing these labor-market frictions, which include search and monitoring costs, among other unobserved costs, the demand for labor by an individual household will increase, especially for the more shadow wage responsive farm households. This hiring, in turn, is associated with higher levels of revenue both for farms and for NFEs, while increased employment opportunities at either the extensive or intensive margin benefit wage earners, especially the landless poor. Therefore, reducing the barriers to a more efficient labor market should be poverty reducing for both demanders and suppliers of rural labor.

Annex 4A. Tables

TABLE 4A.1: Plot and Community Characteristics for Farms (2001)

	Mean
Plot characteristics	
Hillside (1 = plot is on the hillside)	.217 (.412)
Hilltop (1 = plot is on the top of a hill)	.116 (.32)
Eroded (1 = plot is eroded)	.732 (.443)
Sandy (1 = plot soil is sandy)	.146 (.353)
Pest (1 = plot experienced a pest attack in the last year)	.372 (.483)
Weather (1 = plot experienced a weather shock in the last year)	.516 (.5)
Community characteristics	
Transport cost	10366 (14,015)
Zone rouge (1 = yes)	.154 (.361)
Access to broadcast media (1 = yes)	.431 (.495)
Access to finance (1 = yes)	.0704 (.256)
N	7,671

Note: Standard deviation in parentheses. Transport costs reflect cost of transporting 50 kilograms of rice to nearest urban center in the rainy season.

TABLE 4A.2: Agricultural Inputs by Crop Type (Means, 2001)

	Rice	Non-rice food crops	Export crops
Area (*ares* = .01 ha)	68.89 (114.6)	46.30*** (101.8)	66.15 (89.87)
Family labor (days)	51.56 (79.56)	25.76*** (49.57)	43.94** (70.40)
Hired labor (days)	15.04 (46.67)	7.895*** (40.77)	2.986*** (7.745)
Animal, own (hours)	39.24 (403.6)	17.24*** (216.0)	0.0485*** (0.547)
Tractor, own (hours)	24.75 (436.4)	11.46 (279.3)	0.0544*** (0.515)
NPK fertilizer (kg)	2.831 (46.29)	1.928 (66.28)	0*** (0)
TLU	2.439 (6.945)	3.599*** (10.80)	1.054*** (2.221)
Equipment value, log	3.362 (1.318)	3.272*** (1.347)	3.316 (1.170)
N	3,479	2,814	423

Note: Standard deviation in parentheses. TLU = tropical livestock unit.
***, **, *Significantly different from rice at 1%, 5%, 10% levels respectively

TABLE 4A.3: NFE Summary Statistics

	2001 mean	2005 mean	2010 mean
Enterprise has had actual activity in the last year (1 = yes)	0.979 (0.143)	0.961[a] (0.209)	0.953[a,b] (0.212)
Wage paid to household members (MGA/day)	—	8.751 (161.2)	10.15 (97.81)
Wage paid to hired workers (MGA/day)	—	52.51 (463.3)	59.54 (484.9)
Received financial aid	0.011 (0.002)	0.013 (0.002)	0.018[a,b] (0.002)
Number of household employees	1.231 (0.91)	1.486[a] (0.931)	1.634[a,b] (1.114)
Number of hired employees	0.356 (1.645)	0.345 (1.668)	0.451[a] (1.442)
Value of equipment (10,000 MGA)	111.8 (6965)	187.5[a] (2,905)	264.8[a] (2,945)
Years in operation	6.166 (9.495)	6.369 (7.468)	8.285[a,b] (8.832)
Number of enterprises operated by a household	1.321 (0.536)	1.211[a] (0.448)	1.118[a,b] (0.368)
Agriculture enterprise (1 = yes)	0.0523 (0.223)	0.042 (0.201)	0.28[a,b] (0.449)
Manufacturing enterprise (1 = yes)	0.154 (0.361)	0.024[a] (0.153)	0.0422[a] (0.201)
Trade enterprise (1 = yes)	0.174 (0.379)	0.0723[a] (0.259)	0.342[a,b] (0.474)
Services enterprise (1 = yes)	0.481 (0.500)	0.52[a] (0.500)	0.336[a,b] (0.472)
Monthly wages paid[3]	30.00 (189.75)	—	—
N	1,568	3,333	5,783

Note: Standard deviation in parentheses.
Financial aid is the amount, in MGA, that the enterprise "benefitted from" over the past 12 months. Possible sources (from codes on the survey) include microfinance institutions, help from parents or friends, or government grants.
Wages are not differentiated between household and nonhousehold labor in 2001.
[a]Significantly different from 2001. [b]Significantly different from 2005.

TABLE 4A.4: Farm Production Function Estimates (2001)

	Gross revenue
Inputs	
Total labor (days)	0.0190*** (0.0031)
Area (*ares*)	–0.0525*** (0.0024)
NPK (kilograms)	0.0230** (0.0095)5)
Tractor, own (hours)	0.0007 (0.0052)
Animal traction, own (hours)	0.0032 (0.0033)5)
Total labor^2	–0.0004*** (0.0002)
Total labor x area	0.0002 (0.0002)
Total labor x NPK	–0.0019 (0.0013)
Total labor x tractor	0.0002 (0.0011)
Total labor x animal traction	–0.0006* (0.0004)

(continued)

TABLE 4A.4: Farm Production Function Estimates (2001) *(continued)*

	Gross revenue
Area^2	0.0013*** (0.00008)
Area x NPK	0.0008 (0.0009)
Area x tractor	0.00005 (0.0005)
Area x animal traction	0.0005 (0.0003)
NPK^2	–0.0003 (0.0002)
NPK x tractor	–0.0032 (0.0103)
NPK x animal traction	0.0031* (0.0016)
Tractor^2	–0.0001 (0.0001)
Tractor x animal traction	0.0003 (0.0002)
Animal traction^2	-0.000002 (0.00003)
Equipment value	0.000003 (0.00001)
TLU	–0.0004 (0.0005)
Household head education (none/some primary omitted)	
Completed primary school	0.0463* (0.0262)
Some high school	0.0085 (0.0159)
Completed high school	0.0161 (0.0455)
Post high school	–0.0208* (0.0113)
Plot characteristics (1 = yes)	
Plot is eroded	0.0082 (0.0100)
Plot is sandy	–0.0109 (0.0104)
Pest attack	–0.0257*** (0.0087)
Weather shock	–0.0077 (0.0085)
Household owns plot	0.0543*** (0.0142)
Hillside plot	–0.0041 (0.0107)
Hilltop	–0.0020 (0.0138)
Community characteristics	
Transport cost	-0.0000004 (0.0000004)
Access to broadcast media	–0.0180 (0.0144)
Access to finance	0.0206 (0.0215)
Zone rouge	–0.0092 (0.0135)
Electricity available	–0.0260* (0.0144)
Irrigation available	0.0152 (0.0107)
Household (Fixed) Effect	No
Constant	3.548*** (0.0389)
N	3,133

Note: Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1.

TABLE 4A.5: NFE Net Revenue Function Estimates (2001, 2005, and 2010)

	2001 Revenue[1]	2005 Revenue[1]	2010 Revenue[1]
Inputs			
Total labor (person-months)	0.0584*** (0.0145)	–0.206** (0.0977)	0.0805*** (0.0259)
Equipment value, ln (MGA)	–0.137*** (0.0160)	-0.0537 (0.0340)	–0.0999*** (0.0137)
Total labor^2	0.0127*** (0.00395)	0.257*** (0.0573)	0.0458** (0.0225)
Total labor x equipment value	–0.0136*** (0.00418)	–0.0380 (0.0313)	-0.00400 (0.00692)
Equipment value^2	0.0480*** (0.00446)	0.0662*** (0.0111)	0.0731*** (0.00603)

(continued)

TABLE 4A.5: NFE Net Revenue Function Estimates (2001, 2005, and 2010) *(continued)*

	2001 Revenue[1]	2005 Revenue[1]	2010 Revenue[1]
NFE characteristics			
Years in operation	0.0002 –0.0004	0.00169* (0.000895)	0.002*** (0.0004)
Received financial aid (1 = yes)	–0.0236 (0.0206)	0.0170 (0.0517)	0.0760*** (0.0269)
Agriculture	–0.0152 (0.0237)	0.117*** (0.0296)	0.00436 (0.00315)
Manufacturing	–0.0302** (0.0145)	0.0796* (0.0434)	0.0163* (0.00851)
Trade	–0.0307** (0.0125)	0.116*** (0.0272)	0.0276*** (0.00432)
Services	0.00993 (0.0107)	0.106*** (0.0151)	—
TLU	0.0006 –0.0004	-0.00796 (0.00914)	—
Land area	–0.0001** (0.00005)	–0.0159 (0.0295)	—
Household head education (none/some primary omitted)			
Completed primary	0.0283 (0.0178)	0.00914 (0.0158)	0.0950*** (0.0192)
Some high school	0.0381*** (0.0088)	0.0766*** (0.0184)	0.108*** (0.0101)
Completed high school	0.0621*** (0.0129)	0.0430 (0.0341)	0.107*** (0.0234)
Post high school	0.00627 (0.0096)	—	–0.0172* (0.00905)
Community characteristics			
Access to finance	0.0109 (0.0281)	0.0710*** (0.0219)	–0.0280* (0.0165)
Access to broadcast media	–0.0229 (0.0148)	0.0720 (0.0461)	–0.00819 (0.0207)
Zone rouge	0.00325 (0.0195)	–0.077*** (0.0224)	–0.0224 (0.0292)
Transport cost	–0.0001*** (0.00005)	–0.000004* (0.000002)	–0.000005 (0.000003)
Electricity available	—	0.00133 (0.0207)	0.0444*** (0.0150)
Irrigation available	—	–0.00320 (0.0187)	0.0277** (0.0135)
Household (fixed)			
Effect	No	No	No
Constant	4.078*** (0.221)	2.386*** (0.0756)	3.145*** (0.0607)
Observations	1,510	1,382	5,373

Note: NFE = nonfarm enterprise. Standard errors in parentheses.***p<0.01, **p<0.05, *p<0.1

TABLE 4A.6: Estimated Elasticities of Revenue and Marginal Revenue Product of Labor

Elasticities	NFE 2001	NFE 2005	NFE 2010	Farm 2001
Total labor	0.018 (0.006)	0.0081[a] (0.002)	0.009[a] (0.0015)	0.301 (0.801)
Land area	—	—	—	–0.072 (0.292)
NPK	—	—	—	0.001 (0.0039)
NFE equipment	–0.078 (0.011)	–0.016 (0.0121)	–0.0284 (0.0041)	
Marginal revenue product				
Total labor (MGA)	60,647.2 (4001.2)	58,321.6 (1147.7)	27,231.7[a] (80876.5)	7,689.131 (1302.234)
Land (MGA)	—	—	—	–30,797.18 (432,491.5)
NPK (MGA)	—	—	—	828.488 (1987.6)
NFE equipment	–19,651.2 (7332.12)	–14,002.6[a] (9768.81)	–7454.57[a] (2618.7)	
Allocative inefficiency, means	–2.331 (.226)	–2.567[a] (.291)	–1.918[a] (.147)	–1.578 (.070)
Allocative inefficiency, medians	–1.502	–2.065[a]	–1.872[a]	–1.573

Note: Standard deviations in parentheses. NPK = fertilizer. NFE = nonfarm enterprise.
[a] Significantly different from 2001 value.

TABLE 4A.7: Estimation of Naïve Allocative Inefficiency Factor (AIF) in Labor Hiring Decision

	Farms 2001 AIF	NFEs 2001 AIF	NFEs 2005 AIF	NFEs 2010 AIF
Number of working-aged men (ln)	0.0709 (0.0702)	0.789 (0.515)	–0.231 (0.279)	0.155 (0.170)
Number of working-aged women (ln)	–0.0255 (0.0741)	–0.294 (0.495)	–0.0845 (0.324)	–0.190 (0.199)
Number of children (ln)	–0.113 (0.0782)	–0.433 (1.147)	–0.248 (0.445)	0.142 (0.307)
Age of head (ln)	0.00474 (0.104)	–0.458 (0.584)	–0.450 (0.404)	0.114 (0.256)
Head is male	–0.0238 (0.0855)	0.236 (0.519)	–0.350 (0.345)	–0.287 (0.226)
Head is a migrant	0.0514 (0.125)	–0.146 (0.393)	–0.138 (0.211)	—
Household head education (some primary/none omitted)				
Completed primary	0.251* (0.148)	0.495 (0.822)	0.483 (0.377)	–0.276 (0.357)
Some high school	0.239* (0.128)	0.947** (0.406)	0.461 (0.369)	–0.0623 (0.177)
Completed high school	0.139 (0.516)	1.048** (0.500)	—	0.118 (0.283)

(continued)

TABLE 4A.7: Estimation of Naïve Allocative Inefficiency Factor (AIF) in Labor Hiring Decision *(continued)*

	Farms 2001 AIF	NFEs 2001 AIF	NFEs 2005 AIF	NFEs 2010 AIF
Post high school	–0.0813 (0.0643)	0.782 (0.496)	0.590 (0.413)	–0.117 (0.207)
Equipment value, ln	0.0194 (0.0207)	0.103*** (0.0304)	–0.0134 (0.0370)	–0.0555** (0.0244)
TLUs	0.00390 (0.00399)			
Total land (ln)	0.0846*** (0.0267)			
Distance from plot to village (minutes walking)	0.000455 (0.00130)			
Rice	–0.238*** (0.0621)			
Export crop	–0.472*** (0.122)			
Manufacturing enterprise		0.235 (0.720)	0.116 (0.508)	0.440* (0.245)
Trade enterprise		–0.186 (0.448)	0.0454 (0.325)	–0.386** (0.177)
Services enterprise		0.141 (0.392)	–0.655*** (0.234)	0.672*** (0.205)
Number of enterprises, per household		0.145 (0.343)	0.433** (0.193)	–0.264 (0.170)
NFE has actual activity[1]		–0.150 (0.849)	0.764 (0.637)	–0.130 (0.385)
Years in operation (ln)		0.202 (0.157)	–0.0306 (0.110)	0.0500 (0.0682)
Constant	15.53*** (0.645)	12.15*** (2.593)	–4.713*** (1.724)	–1.379 (1.011)
District Fixed Effects	Yes	Yes	Yes	Yes
n	432	204	202	427
R-squared	0.644	0.220	0.119	0.111

Note: Standard errors in parentheses. NFE = nonfarm enterprise. ln=natural logarithm. *** p<0.01, ** p<0.05, * p<0.1.

TABLE 4A.8: Estimated Shadow Wages and Observed Wages (MGA per Day)

	Farm 2001 mean	NFE 2001 mean	NFE 2005 mean	NFE 2010 mean
Shadow wages				
Nonhiring enterprise	7434.72 (2864.544)	7478.84 (25316.15)	3233.421[a] (5429.388)	3869.321[a] (9386.16)
Hiring enterprise	7512.422** (955.843)	10287.15*** (16112.61)	6945.05[a,]** (12015.4)	8008.15[a,]** (1817.816)
Observed wages	5652.614††† (927.656)	6867.779††† (32226.84)	2667.61[a,]††† (5445.34)	02049.57[a,]††† (14927.18)

Note: Standard deviation in parentheses. NFE = nonfarm enterprise. ***, **, *Significantly different from hiring at 1%, 5%, 10% levels respectively. †††,††,†Significantly different from MRP_L at 1%, 5%, and 10% levels respectively. [a]Significantly different from 2001.

TABLE 4A.9: Farms: Estimated Demand for Labor

	Farm 2001		
	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)
Shadow wage (log (ln) MGA)	–1.599*** (0.107)	–1.324*** (0.085)	–0.1024*** (0.01463)
Number of working-age men (ln)	–0.0138 (0.104)	0.0283 (0.0997)	0.0022 (0.00773)
Number of working-age women (ln)	0.0910 (0.0906)	0.0818 (0.0791)	0.0063 (0.00626)
Number of children (ln)	–0.136 (0.107)	–0.117 (0.092)	–0.0091 (0.00749)
Age of head (ln)	–0.164 (0.157)	–0.107 (0.1361)	–0.0083 (0.01077)
Head is male	–0.0560 (0.118)	–0.0561 (0.1001)	–0.0041 (0.00705)
Household head education (some primary/none omitted)			
Completed primary	0.376 (0.263)	0.2814 (0.231)	0.01602 (0.0100)
Some high school	0.340* (0.198)	0.3230* (0.1923)	0.0182** (0.0078)
Completed high school	–1.312 (0.895)	–0.7562 (0.0877)	–0.1561 (0.2531)
Post-high school	–0.157 (0.0960)	0 (0)	–0.0103 (0.0073)
Transport cost	–1.83e-06 (4.13e-06)	0.1385 (0.0896)	0 (0)
Access to broadcast media	0.227** (0.108)	0.1028 (0.1317)	0.0106 (0.0070)
Access to finance	0.143 (0.147)	0.111 (0.1198)	0.0072 (0.0084)
Zone rouge	0.657*** (0.121)	0.6276*** (0.1198)	0.0325*** (0.0068)
Equipment value (ln)	–0.0139 (0.0274)	–0.0044 (0.0262)	–0.0003 (0.0020)
TLU	0.0196*** (0.00709)	0.1358*** (0.0295)	0.0013*** (0.0005)
Total land (ln)	0.162*** (0.0339)	–0.444*** (0.0785)	0.0105*** (0.0023)
Rice	–0.531*** (0.0933)	–0.808*** (0.1301)	–0.0434*** (0.01061)
Export crop	–1.213*** (0.175)	0.2814 (0.24)	–0.0374 (0.03169)
Province FE	Yes	Yes	Yes
Constant	20.69*** (1.484)		
σ	1.032*** (0.0334)		
n	2,390	2,390	2,390

Note: Standard errors in parentheses. *Unconditional* refers to the unconditional expectation of the observed dependent variable
*** p<0.01, ** p<0.05, * p<0.1.

TABLE 4A.10: NFEs—Estimated Demand for Labor

	2001			2005			2010		
	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)
Shadow wage (n) MGA)	–0.270*** (0.0229)	–0.1040*** (0.01758)	–0.129*** (0.0222)	–0.0555*** (0.0149)	–0.0267 (0.1481)	–0.0148 (0.1748)	–0.0190** (0.00948)	–0.0147* (0.0078)	–0.0039* (0.0021)
Men (ln)	0.207*** (0.022)	0.2140*** (0.0311)	0.265*** (0.0393)	0.407*** (0.155)	–0.0686 (0.3770)	–0.0374 (0.4460)	0.282*** (0.0265)	0.2320*** (0.0214)	0.0613*** (0.0071)
Women (ln)	0.151** (0.085)	0.1330*** (0.0349)	0.165*** (0.0432)	0.620*** (0.144)	0.0400 (0.2333)	0.0218 (0.2653)	0.165*** (0.0286)	0.1361*** (0.0235)	0.0360*** (0.0069)
Children (ln)	0.458 (0.347)	–0.0866 (0.0612)	–0.1074 (0.0758)	0.271 (0.201)	–0.1681 (1.1316)	–0.0917 (1.144)	0.263*** (0.0384)	0.2165*** (0.0309)	0.0572*** (0.0089)
Age of head (ln)	0.341* (0.186)	–0.0454 (0.0317)	–0.0563 (0.0395)	0.148 (0.121)	–0.0069 (0.0893)	–0.0038 (0.0656)	0.000360 (0.0342)	0.0003 (0.0284)	0.0001 (0.0075)
Head is male	–0.556** (0.143)	–0.0792*** (0.0267)	–0.098*** (0.0331)	0.270*** (0.101)	0.1395 (0.7534)	0.0705 (0.8994)	–0.0291 (0.0269)	–0.0237 (0.0220)	–0.0060 (0.0053)
Household head education									
Primary school	0.307*** (0.054)	0.2036*** (0.0668)	0.219*** (0.0607)	0.0733 (0.0919)	0.0500 (0.2843)	0.0273 (0.3290)	–0.0216 (0.0577)	–0.0179 (0.0472)	–0.0050 (0.0137)
Some high school	0.270* (0.131)	0.1306*** (0.0332)	0.155*** (0.0377)	0.178 (0.123)	0.0398 (0.2305)	0.0221 (0.2617)	–0.0635** (0.0267)	–0.0519** 0.0208	–0.0151** (0.0068)
Completed high school	0.248* (0.139)	0.2849*** (0.0605)	0.293*** (0.0486)	0.282 (0.248)	0.0850 (0.4967)	0.0496 (0.5488)	–0.138* (0.0737)	–0.1086** (0.0548)	–0.0388 (0.0260)
High school+	0.115 (0.144)	—	—	—	—	—	0.00800 (0.0202)	—	—
Transport cost, ln	0.000003 (0.000006)	–0.0008 (0.0072)	–0.0011 (0.0089)	–0.00002 (0.00002)	0.0672 (0.3714)	0.0366 (0.4343)	0.0143** (0.00609)	0.0118** (0.0049)	0.0031** (0.0013)
Broadcast media	–0.181 (0.152)	0.0485 (0.0378)	0.0587 (0.0445)	0.558*** (0.210)	0.6683 (3.8151)	0.3644 (4.392)	–0.0595 (0.0497)	–0.0497 (0.0440)	–0.0116 (0.0091)
Access to finance	–0.539** (0.154)	–0.1746*** (0.0535)	–0.226*** (0.0678)	–0.0598 (0.441)	–0.0046 (0.1227)	–0.0024 (0.0715)	0.0171 (0.0374)	0.0142 (0.0317)	0.0036 (0.0078)
Access to electricity	—	—	—	0.0331 (0.160)	0.0083 (0.113)	.0074 (0.0786)	0.0332 (0.0332)	0.0271 (0.0277)	0.0074 (0.0079)

(continued)

TABLE 4A.10: NFEs—Estimated Demand for Labor *(continued)*

	2001			2005			2010		
	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)	(1) Unconditional	(2) Conditional (intensive margin)	(3) Probability (extensive margin)
Access to irrigation	—	—	—	0.268* (0.138)	0.0418 (0.540)	0.0374 (0.3517)	0.00600 (0.0319)	0.0049 (0.0241)	0.0013 (0.0064)
Zone rouge	0.240*** (0.022)	0.1165** (0.0574)	0.1346** (0.0615)	–0.0958 (0.160)	–0.0557 (0.3235)	–0.0304 (0.3669)	–0.0281 (0.0878)	–0.0229 (0.0738)	–0.0064 (0.0220)
Equipment value, ln	0.080*** (0.008)	0.0163*** (0.0020)	0.020*** (0.0026)	0.118*** (0.0262)	0.0522 (0.2862)	0.0285 (0.3403)	–0.0167*** (0.00501)	–0.0139 (0.0041)	–0.0037*** (0.0011)
TLU	0.142** (0.001)	0.0050*** (0.0015)	0.006*** 0.0018	–0.0249 (0.0518)	0.0118 (6.520)	0.0064 (3.578)	–0.00137 (0.00203)	–0.0011 (0.0017)	–0.0003 (0.0005)
Land area	–0.0001 (0.001)	–0.0003 (0.0002)	–0.0004 (0.0003)	–0.0495 (0.101)	–0.1275 (48.818)	–0.0695 (26.184)	0.0001 (0.000959)	0.00001 (0.00001)	0.0000 (0.000)
Agriculture	0.118 (0.240)	–0.0103 (0.0630)	–0.0128 (0.0790)	–0.184 (0.224)	—	—	–0.0778*** (0.0232)	–0.063*** (0.0185)	–0.0182*** 0.0060
Manufacturing	–0.0810 (0.194)	0.0786* (0.0446)	0.0940* (0.0510)	0.894** (0.371)	—	—	–0.466*** (0.0731)	–0.321*** (0.0388)	–0.2130*** (0.0523)
Trade	–0.223 (0.168)	–0.0304 (0.0357)	–0.0380 (0.0452)	1.331*** (0.319)	—	—	–0.0486** (0.0229)	–0.040** (0.0188)	–0.0110** (0.0055)
Services	0.056 (0.143)	0.0269 (0.0325)	0.0333 (0.0402)	0.409*** (0.0947)	—	—	—	—	—
Years in operation	–0.003 (0.003)	0.0021* (0.0011)	0.0026* (0.0014)	0.00838 (0.00525)	0.0025 (0.0142)	0.0014 (0.0165)	0.00382*** (0.00119)	0.0031*** (0.0009)	0.0008*** (0.0003)
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.199*** (0.761)			0.537 (0.429)			0.784*** (0.141)		
Sigma	0.997*** (0.0312)			0.413*** (0.0201)			0.431*** (0.0103)		
n	1369	1369	1369	1589	1589	1589	2,462	2,462	2,462

Note: Standard errors in parentheses. NFE = nonfarm enterprise. *Unconditional* refers to the unconditional expectation of the observed dependent variable. TLU = tropical livestock unit. Ln=natural logarithm.
***p<0.01, **p<0.05, *p<0.1.

Annex 4B. Summary Statistics by Gender

TABLE 4B.1: Agriculture: Plot and Community Characteristics: Male- and Female-Headed Households

	Male		Female	
	mean	n	mean	n
Plot Characteristics				
Hillside	0.173***	8210	0.132	1821
	(0.378)		(0.339)	
Hilltop	0.0895	8210	0.0846	1821
	(0.286)		(0.278)	
Eroded	0.576***	8210	0.487	1821
	(0.494)		(0.5)	
Sandy	0.115**	8210	0.0956	1821
	(0.32)		(0.294)	
Pest attack	0.285***	8210	0.235	1821
	(0.452)		(0.424)	
Weather shock	0.401***	8210	0.344	1821
	(0.49)		(0.475)	
Community Characteristics				
Average transport cost	11125***	5576	12835	993
	(12916)		(13366)	
"Zone rouge" indicator	0.178	5530	0.197	985
	(0.383)		(0.398)	
Access to broadcast media	0.518***	5530	0.446	985
	(0.5)		(0.497)	
Access to finance	0.073	6205	0.0757	1150
	(0.26)		(0.265)	

Note: Standard deviation in parenthesis; ***, **, *Significantly different from female at 1%, 5%, 10% respectively.

As shown in table 4B.1, male-headed households are more likely to operate hillside plots, eroded plots, or sandy plots. Rather than this indicating some sort of preferential treatment for female-headed households, this may reflect the greater likelihood that a male-headed household will operate any plot, regardless of its quality or position. Similarly, male-headed households experience both weather and pest shocks more frequently than female-headed households. It does seem, however, that women operate plots in communities that are more remote: the transport costs are significantly lower for male-headed households, and they are also more likely to have access to broadcast media.

TABLE 4B.2: Agriculture—Summary Statistics: Male- and Female-Headed Households

	Male		Female	
	mean	n	mean	n
Household owns plot	91.40% (0.28)	8210	90.20% (0.298)	1821
Area (hectares)	59.83 (107.3)	5642	56.18 (115.1)	1039
Family labor (days)	41.15 (71.36)	5673	35.42 (54.91)	1043
Hired labor (days)	11.49 (45.5)	5674	10.17 (24.78)	1043
Animal traction, own (hours)	31.61** (349.9)	5674	5.48 (58.01)	1043
Animal traction, rented (hours)	299.4** (4440)	5673	635.00 (6541)	1043
Tractor, own (hours)	20.83* (394.2)	5674	0.16 (0.867)	1043
Tractor, rented (hours)	523.1** (3726)	5674	405.50 (2717)	1043
Times weeding	1.77 (6.867)	5674	1.79 (4.372)	1043
NPK (kilogram)	2.58 (59.07)	5674	0.64 (3.987)	1043
Urea (kilogram)	6.17 (106.1)	5674	7.02 (70.61)	1043
Organic fertilizer (MGA)	9065*** (15919)	1211	4,831.00 (9840)	172
Pesticide (MGA)	2030* (21387)	5674	864.20 (8763)	1043
Agricultural equipment (MGA)	432,807.00 (8423000)	7698	218,879.00 (3744000)	1659
Revenue per acre	529,484.00 (11760000)	3635	132,433.00 (484610)	684
Revenue per month of labor	11,930,000.00 (279000000)	3560	2,028,000.00 (5372000)	681
Revenue per value of equipment	118.1*** (1638)	3309	1,518.00 (19630)	639

Note: Standard deviation in parentheses. ***, **, *Significantly different from female at 1%, 5%, 10%, respectively.

TABLE 4B.3: NFEs—Summary Statistics: Male- and Female-Headed Households

	2001		2005		2010	
	male mean	female mean	male mean	female mean	male mean	female mean
Enterprise has had actual activity in the last year	0.979 (0.145)	0.981 (0.136)	0.956** (0.223)	0.974 (0.161)	0.952 (0.213)	0.954 (0.209)
Salary paid to household members	—	—	10.74* (184.5)	2.933 (48.45)	10.04 (97.16)	10.58 (100.4)
Daily salary paid to hired workers (100s MGA)	—	—	63.58*** (512.4)	19.97 (269.7)	67.18*** (526.4)	28.66 (254.8)
Number of household employees	1.263** (0.922)	1.129 (0.866)	1.514*** (0.924)	1.405*** (0.945)	1.664 (1.145)	1.511 (0.972)
Number of hired employees	0.408*** (1.788)	0.194 (1.071)	0.405*** (1.878)	0.165*** (0.739)	0.496 (4.907)	0.265 (1.272)
Value of equipment	1353000*** (7943000)	376537 (1403000)	229.5*** (2,411)	63.78 (389.5)	312.2*** (3,271)	73.11 (679.3)
Years in operation	6.153 (9.783)	6.206 (8.544)	6.153*** (7.094)	7.004 (8.447)	8.035*** (8.512)	9.301 (9.966)
Number of enterprises operated by a household	1.35*** (0.546)	1.233 (0.493)	1.229*** (0.461)	1.159 (0.403)	1.126*** (0.377)	1.086 (0.324)
Agriculture enterprise	0.0504 (0.219)	0.0582 (0.234)	0.0486*** (0.215)	0.0225 (0.149)	0.285** (0.452)	0.257 (0.437)
Manufacturing enterprise	0.161 (0.368)	0.132 (0.339)	0.0309*** (0.173)	0.00356 (0.0596)	0.0504*** (0.219)	0.00876 (0.0932)
Trade enterprise	0.176 (0.381)	0.169 (0.376)	0.0851*** (0.279)	0.0344 (0.182)	0.326*** (0.469)	0.406 (0.491)
Services enterprise	0.468* (0.499)	0.521 (0.500)	0.507** (0.500)	0.556 (0.497)	0.338 (0.473)	0.328 (0.47)
Monthly wages paid	400528 (2147000)	232474 (2644000)	—	—	—	—
N	1,190	378	2,490	843	4,641	1,142

Note: Standard deviation in parentheses. ***, **, *Significantly different from female at 1%, 5%, 10%, respectively.

NOTES

1. Further, the failures they find are widespread and structural in nature and unrelated to household characteristics, including gender of the household head and remoteness.
2. Community-level access to services including irrigation and electricity were not significant in our estimations of the marginal revenue product of labor on-farm in that year, and therefore we found no evidence of an effect on labor demand of these types of infrastructure in these data. However, applying labor and fertilizer (NPK) both contributed positively to farm plot-level revenue. The effect of individual (within-household) plot size on revenues was not significant, possibly because there is an inverse relationship between land quality and plot size in Madagascar.
3. This poverty rate is for 2012 based on the World Bank's extreme poverty line of $1.90 per day per person (2011 U.S. purchasing power parity dollars).
4. Unfortunately, the data do not report the activities of the laborers, whether they are family or not. While there is no significant difference between the quantities of labor, it is entirely possible that the level of expenditure on hired labor does differ, and that men and women hire labor to perform very different tasks.
5. In the production function estimation, hired labor and family labor are aggregated together to overcome issues of dimensionality in later steps. They are presented here separately to highlight differences in the amount used of each.
6. Summary statistics are calculated using population weights.
7. We use the primal approach to estimation, as the dual approach, which requires input price data and (ex ante expected) output prices, is infeasible. Of course, the problem of unobserved wages (that is, the price of labor) is exactly the problem this paper's method is designed to overcome.
8. Gross revenue is not used as the dependent variable for the farm and nonfarm production function because gross revenue is not recorded for NFEs on the survey instrument. Instead, the survey asks for revenue but defines it as total revenue minus certain expenses. Due to the difficulty for respondents of performing that calculation, it is possible respondents actually recorded gross revenue for this question. Summary statistics, which show no negative values in response to this question, support this possibility.
9. Financial aid is the amount, in MGA, that the enterprise "benefitted from" over the past 12 months. Possible sources (from codes on the survey) include microfinance institutions, help from parents or friends, or government grants.

10. This is an indicator variable where *1* indicates that the respondent answered "yes" to the question "Did this enterprise have actual business activity in the last 12 months?"
11. Note that we generalize this functional form further by allowing inputs to enter the revenue function as polynomials (with a squared term).
12. It is labeled"naïve" because the observed inefficiency does not imply an error, or waste, by the decision maker; it merely indicates divergence between w and $M\hat{R}P_L$.
13. Bootstrapping can also be sued where the distribution of error terms is not known. See Horowitz (2001).
14. It is possible to calculate the decomposition manually, but is most easily and efficiently achieved using the Stata command "margins," which is what we used here.
15. There are also positive externalities from a labor demand perspective for the physical insecurity indicator, indicating that not all increased labor demand reflects improvements in overall well-being.

REFERENCES

Barrett and Dillon. 2016. "Agricultural factor markets in Sub-Saharan Africa: An updated view with formal tests for market failure." *Food Policy.* Article in Press.

Barrett, C. B., T. Reardon, and P. Webb. 2001. "Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications." *Food Policy* 26 (4): 315–31.

Barrett, C. B., S. M. Sherlund, and A. A. Adesina. 2008. "Shadow Wages, Allocative Inefficiency, and Labor Supply in Smallholder Agriculture." *Agricultural Economics*, 38 (1): 21–34.

Behrman, J. R. 1999. "Labor Markets in Developing Countries." O. Ashenfelter, D. Card (Eds.), *Handbook of Labor Economics*, vol. 3B: 2859–2939.

Belghith, N., P. Randriankolona, and T. Osborne. 2016. "Madagascar Poverty and Inequality Update: Recent Trends in Welfare, Employment, and Vulnerability." Washington, DC: World Bank.

Benjamin, D. 1992. "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models." *Econometrica* 60(2): 287-322.

Chambers, R. G. 1988. *Applied Production Analysis: A Dual Approach*. Cambridge, UK: Cambridge University Press.

Hammermesh, D. S. 1996. *Labor Demand*. Princeton, NJ: Princeton University Press.

Horowitz, J. L. 2001. "The Bootstrap." *Handbook of Econometrics*, *5*, 3159–3228.

IFAD (International Fund for Agricultural Development). 2011. *Rural Poverty Report.* www.ifad.org/rpr2011.

Jacoby, H. G. 1993. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *Review of Economic Studies* 60(4): 903–21.

McDonald, J. F., and R. A. Moffitt. 1980. "The Uses of Tobit Analysis." *Review of Economics and Statistics* 62(2): 318–21.

Otsuka, K., and T. Yamano. 2006. "The Role of Rural Labor Markets in Poverty Reduction: Evidence from Asia and East Africa." Background paper for World Bank. 2008. *World Development Report 2008: Agriculture for Development*. Washington, DC: 2008.

Randrianarisoa, J. -C., C. B. Barrett, and D. C. Stifel. 2009. *The Demand for Hired Labor in Rural Madagascar*. Washington, DC: USAID. http://www.cfnpp.cornell.edu/images/wp234.pdf.

Skoufias, E. 1994. "Using Shadow Wages to Estimate Labor Supply of Agricultural Households." *American Journal of Agricultural Economics* 76 (2): 215–27.

Stifel, D., M. Fafchamps, and B. Minten. 2011. "Taboos, Agriculture and Poverty." *Journal of Development Studies* 47 (10): 1455–1481.

Stifel, D., F. H. Rakotomanana, and E. Celada. 2007. *Assessing Labor Market Conditions in Madagascar, 2001–2005*. Washington, DC: World Bank.

Thiebaud, A., T. Osborne, and N. Belghith. 2016. "Isolation, Crisis, and Vulnerability: A Decomposition Analysis of Inequality and Deepening Poverty in Madagascar (2005–2010). Washington, DC: World Bank.

CHAPTER

5

Transactions Costs, Poverty, and Low Productivity Traps: Evidence from Madagascar's Informal Microenterprise Sector

Chuqiao Bi and Theresa Osborne*
June 2016
*AFR Poverty Global Practice, World Bank.
The authors thank Frank Vella for valuable suggestions and guidance.

Summary of Results and Policy Implications

This paper utilizes 2012 data on informal, owner-occupied microenterprises (OOMEs) in Madagascar to assess the potential of these enterprises to achieve higher incomes for their owners and offer remunerative employment to workers. We note first that there is significant overlap in activities for single-worker OOMEs—which employ only their owner—versus others, which employ family and other unpaid as well as paid labor. The sectors range from logging and mining to household, transport services, and manufacturing. We first estimate the returns to capital and labor at varying scales of operation, taking account of individual differences in ability or opportunity that may affect the decision to own such an enterprise rather than obtaining other means of employment in Madagascar's labor markets.

The results show that OOMEs that have only the owner working in them—with no family, other unpaid, or paid workers—have significantly lower returns to capital and to the owner's labor, controlling for worker characteristics. Moreover, we find evidence of profitability increasing with scale, but these potential gains are not exploited. Returns to capital are increasing, and there is significant underemployment of workers relative to the profit-maximizing level. The returns to capital for single-worker OOMEs, which comprise 70 percent of OOMEs, average approximately 12 percent per annum (nominal) and are just one-third those of multi-worker OOMEs. These returns fall below prevailing lending rates in the country at the time. The wage penalty, estimated as the average returns to the owner's time, is approximately 60 percent of the mean wage in the broader labor market. Owners of multi-worker OOMEs, however, earn a premium of approximately 68 percent of the mean wage, controlling for individual characteristics.

To derive policy implications requires that we attempt to explain these findings. OOMEs perceive a lack of demand for as their most immediate constraint, and the size of the market certainly plays a major role in reducing enterprise profits. Nonetheless, it would still be more efficient for there to be fewer, larger enterprises serving the same level of demand. OOMEs could in principle increase their profitability by expanding a small amount at a time, reinvesting growing profits and growing to a more efficient and profitable scale. Such a market restructuring would expand incomes and thereby markets, given the importance of the OOME sector to the overall economy. Yet this process does not take place.

We propose a simple theoretical model to explain the persistence of low-productivity OOMEs. In a dynamic, general equilibrium framework, a combination of conditions is needed. One of these is a lack of entry by larger, more efficient, typically formal firms, which

would provoke a restructuring of the market and draw workers into more remunerative work. In addition, in a poor economy like Madagascar's, the very nature of OOMEs inhibits their growth: First, the marginal utility of consumption for poor entrepreneurs is high, and in the presence of increasing returns, the returns to small incremental investments for the smallest OOMEs are low. Thus, entrepreneurs typically consume all the firm's income. Breaking out of this low-productivity equilibrium (or "trap") would require a more substantial increase in scale than poor households can afford. However, because of the difficulties associated with monitoring the use of firm resources, whether by creditors or potential partners, the transaction costs of expansion through external financing are high. Credit costs more to cover the higher enforcement risks and small loan sizes, and partnerships are much less likely to form precisely because entrepreneurs' level of investible capital is low relative to these transactions costs. Similar issues of information asymmetries and incentives ("transactions costs") likely hinder more optimal levels of labor employment as well.

The policy implications of this paper are tentative but are as follows: First, it would likely have little effect to encourage firms to simply register without taking additional steps to improve the credibility of their financial statements. This and other means to strengthen the information environment would be needed to reduce the transactions costs for potential creditors and partners. In the current context, the benefits of registration appear to be primarily to be able to serve larger customers, but this also depends on entrepreneurial ability, and only so many enterprises can serve these limited markets. The second implication is that while very small loans to the tiniest single-worker OOMEs may help to provide employment for the poor, they would have little impact on overall productivity, employment, and wage growth.

Ultimately, significantly reducing the misallocation of capital and labor in Madagascar's economy would require a steadily growing presence of larger, more formal firms that compete for markets. These firms are essential to creating more productive employment, stimulating greater demand, integrating Madagascar's internal markets, and better accessing export markets. Achieving such a transformation will require alleviation of the constraints that such firms face to entering, growing, and competing on an even playing field, while earning acceptable (risk-adjusted) returns. Therefore, an obvious policy implication is to attempt to alleviate the constraints to formal firm investment. At the same time, some measure of investment and productivity growth appears feasible through the development of OOMEs that have already achieved a certain scale, or for those with demonstrable entrepreneurial skills. Given the presence of increasing returns, such firms could invest more and hire more workers if they had access to financing. Therefore, a fourth policy implication is to intensify efforts to improve the monitoring and enforcement environment for creditors, partners, and associations. In

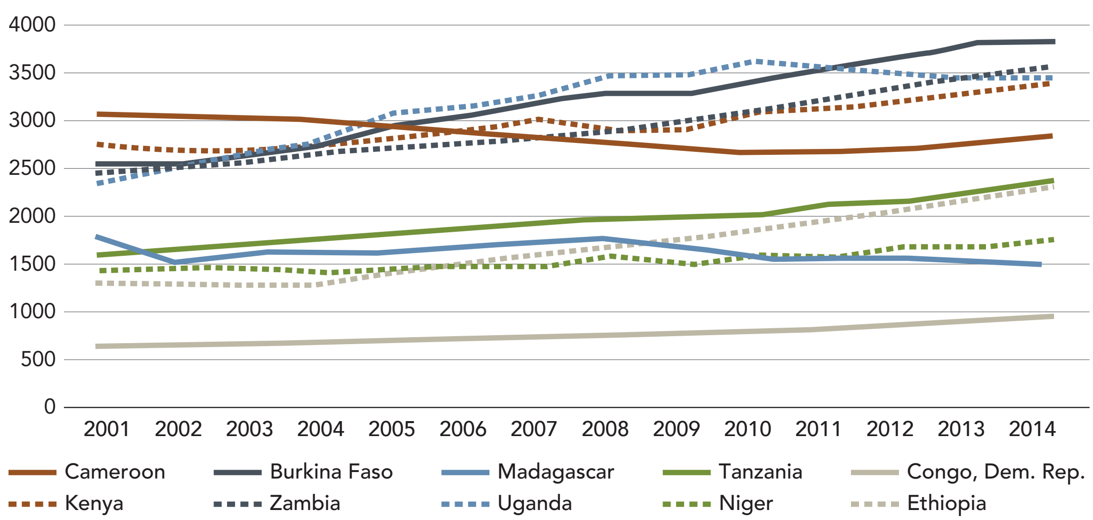
2012, credit registries and bureaus were almost nonexistent in Madagascar, and the strength of legal rights to enforce repayment rated only 2 out of a 10-point scale in Doing Business's *Getting Credit*, although these indicators improved by 2016. Finally, there is an argument for stimulating more rapid evolution of credit markets to serve the higher potential OOMEs. In order to sustain impacts of any such program, it would be important to ensure that (1) the monitoring environment was also improving, (2) that access to such resources was competitive and fairly distributed between and among individual and group enterprises, and (3) that subsidies did not undercut other developments in credit markets. Without a broader expansion of demand for goods and services, as some firms grew, others may be edged out of the market, and thus it is difficult to predict the direction of change in employment levels without better understanding the labor market frictions identified in this paper. Therefore, further investigation into the sources and possible means to address these frictions is needed, as is robust evaluation of the impacts of any credit support programs.

Introduction

Madagascar's high rates of poverty, reaching over 70 percent of the population, are closely associated with its inability to create and sustain productive employment for its workforce. Since 2001, gross domestic product (GDP) per person employed—a broad measure of labor productivity—has fallen in real terms and is now the second lowest in the world (among countries with data) after the Democratic Republic of Congo (see figure 5.1). This is partly because Madagascar's labor force is more concentrated in agriculture, a sector which exhibits particularly poor low labor productivity: 73.2 percent of household heads claim agriculture as their main sector of employment, and 83 percent in the bottom 80 percent of the consumption distribution.[1]

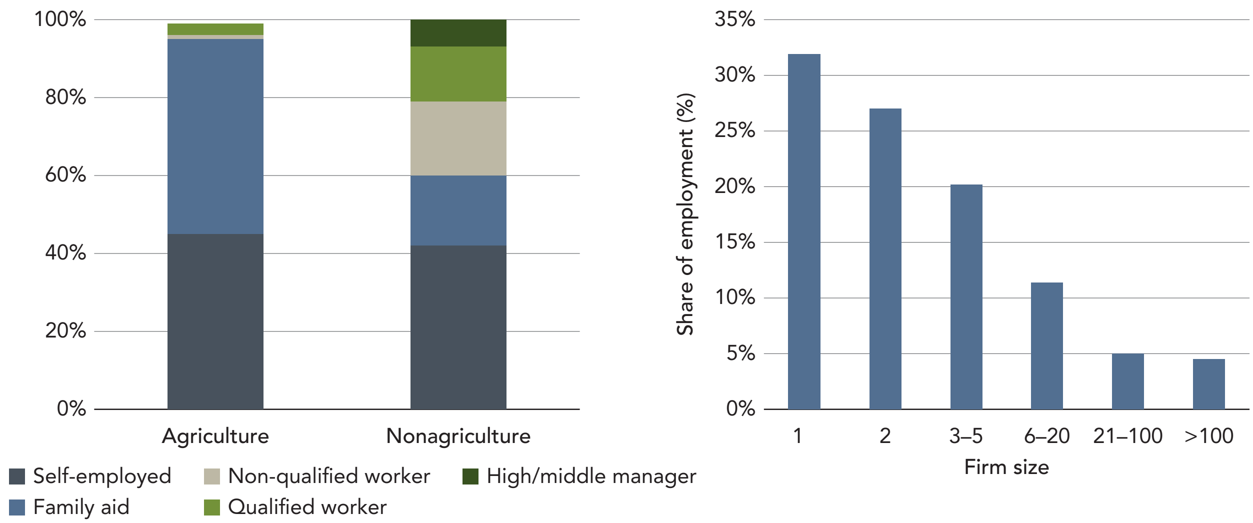
For the 26 percent of the population not primarily employed in agriculture, the vast majority are employed in informal microenterprises. In 2012, over 87 percent of employed workers worked in enterprises with five or fewer workers (ENEMPSI 2012), and 80 percent of nonfarm employees did so (figure 5.2). In contrast, only 4.5 percent of nonagricultural workers worked in establishments with 100 or more workers. In addition, approximately 75 percent of nonfarm jobs in the country were informal (INSTAT 2013), as were approximately 90 percent of new jobs created in 2010. The formal private sector accounts for very little employment, signaling severe constraints on the growth of this sector: Only an estimated 3.9 percent of nonfarm workers were employed by the formal private sector, and only 11 percent of employed workers receive a wage (INSTAT 2013).[2] Although Madagascar is not alone in its prevalence of small, unproductive firms, it appears to be an extreme case.

FIGURE 5.1: GDP per Person Employed, Madagascar and Comparators, 2001–2014 (1990 PPP U.S. Dollars)



Source: World Development Indicators (WDI).
Note: PPP = purchasing power parity.

FIGURE 5.2: Types of Workers and Share of Employment by Firm Size



Source: Sy and Diatta (2014) using data from the LFS 2012.
Note: The panel *b* is restricted to the nonagricultural sector and excludes public administration and public companies.

Although employment in OOMEs is a potential pathway out of poverty for some, a high concentration of labor in such enterprises is also associated with the economy's low productivity.[3] Owner operation and microscale are also strongly associated with informality—the partial or limited use of formal accounting practices, registration, and compliance with regulatory and tax requirements. Employment in these enterprises pays little relative to that in larger, more formal businesses: according to Madagascar's National Institute of Statistics (INSTAT 2013), 84,000 versus 184,000 ariary per month. As a result, 8 out of 10 unemployed workers seek wage employment rather than trying to create a business, and only 8 percent prefer to work for themselves (12 percent are indifferent). Although some existing microentrepreneurs choose to work for themselves or to carry on a traditional family business, according to the 2012 labor force survey (LFS, conducted by INSTAT), 22.7 percent chose their activity because they were unable to find a sufficiently remunerative job. Nonetheless, 46 percent claimed to do so as a way to increase their incomes (table 5.1).

TABLE 5.1: Reasons for Operating an OOME (Percentage by Category)

	Single worker	Multiworker	All
Did not find paid work at large company	6.9	4.8	6.3
Did not find paid work at small business	17.9	12.6	16.4
To get a better income	44.4	49.9	46.0
To be independent	17.0	20.3	17.9
Family tradition	13.8	12.4	13.4

Source: ENEMPSI 2012.

This paper explores the productivity and income implications of there being such a high level of employment of the country's economic resources (capital and labor) in OOMEs. To develop an empirical understanding of the issues, we estimate the contributions to firm profits of capital and labor, as well as how returns change with the scale of operation. In addition, we estimate the effects on profits of key characteristics of the entrepreneur or firm, as well as indicators of levels of competition. We then estimate the effects of this mode of employment on the owner's labor earnings. In all cases, we take into account the possible estimation bias associated with the sorting or selection of workers into microentrepreneurship. Such individuals may be more capable entrepreneurs and/or less employable in the formal sector, may have greater familiarity with business, or may have differential opportunities that are not observed in the data—all sources of "unobserved heterogeneity" which can bias estimates.

We find that the vast majority of OOMEs operate substantially below efficient scale, causing a sizeable inefficiency in the allocation of both capital and labor in the economy. First, our estimates show a significant difference in returns to both capital and labor between OOMEs that utilize only the owner's labor (henceforth "single-worker" OOMEs) and those that also have other workers ("multi-worker" OOMEs), whether those workers are paid or not. Returns to capital by single-worker OOMEs are below the cost of capital in the economy and below those of multi-worker OOMEs, which in contrast earn acceptable returns. Returns to the enterprise head's labor effort in single-worker OOMEs are lower than labor incomes earned elsewhere in the economy. Accounting for observed and unobserved differences in workers, this mode of employment reduces labor earnings by an average of 60,000 ariary (approximately US$30) per month, relative to a mean monthly employment income in urban areas of 101,000 ariary. Yet the returns to the labor of owners of multi-worker OOMEs is on average 68,000 ariary more than it would be in the labor market for equivalent workers.

Second, we find evidence of increasing returns to capital within the population of OOMEs and of increasing returns to scale in much of the relevant range for most OOMEs. As measured, these returns to scale can accrue either through declining average costs or through improved pricing. Those single-worker OOMEs which register their businesses experience a significant increase in profitability, but there is no statistically significant effect of registration on the profits of multi-worker OOMEs, once firm and owner characteristics are taken into account.

In addition, we find evidence of that family history or wealth and gender are important determinants of firm entry and size. Enterprises' current asset levels are significantly related to having a traditional (prior) family business, to the mother's level of education, and to the father's employment history—factors that are likely correlated with the availability of start-up capital. In addition, although the level of enterprise assets is not associated with the gender of the head in the full sample, male ownership increases the profitability of single-worker firms (although there is no evidence that this is the case for multi-worker firms), all else equal. Having more education reduces the likelihood of owning and operating any OOME but positively affects returns once one does. However, these returns are lower than the returns to schooling in the labor market generally. Finally, there is evidence of disadvantageous gender-based sorting into less profitable activities: females are less likely than men to own a multi-worker OOME and more likely to own a single-worker one, and in turn, in the case of single-worker OOMEs, earn lower profits than their male counterparts.

We offer a simple theoretical explanation for the persistent productivity losses observed as a result of suboptimal scale. Our explanation relies on four conditions: First, an important general equilibrium condition outside of our model must prevent the entry of more efficient (typically larger) firms that would compete aggressively with OOMEs and spur Schumpeterian growth, as in this case OOMEs could not compete and would eventually exit (Aghion, Akcigit, and Howitt 2013). We take this condition as given in Madagascar. There is a low reported level of formal firm activity, and explaining this, however critical, is beyond the scope of this paper. Rather, our model focuses on the partial equilibrium constraints to the growth of OOMEs, incorporating the remaining conditions. The second condition is that the initial wealth of entrepreneurs must be low. Third, credit markets must be inadequate to serve a significant fraction of (the most able) entrepreneurs. Third, and relatedly, the costs of mitigating information asymmetries and of enforcing agreements among potential investors must be high. Although there are no data available to provide a formal test of the theory, we provide some descriptive evidence and suggest further avenues of research.

Firm Size Productivity Relationships

The advantages of greater firm size could differ across countries at any time by the countries' areas of comparative advantage, technological readiness, institutional arrangements, policy context, and market access. Nonetheless, a priori, the productivity losses for Madagascar of the prevalence of OOMEs appear likely to be large.

Internationally, countries with a higher GDP per capita tend to have larger firms, fewer microenterprises, and more individuals employed in large firms. See, for example, Poschke (2014). Based upon the most recent comprehensive data on manufacturing firms in 124 countries, average establishment size is strongly correlated with GDP per capita—with an elasticity of 0.26 (Bento and Restuccia 2014). Similarly, using data from 47 developing countries, Ayyagari, Demirgüç-Kunt, and Maksimovic (2011) find that large firms are more innovative and productive than small firms.[4]

Moreover, studies of micro- and small enterprises suggest that informal microenterprises possess little growth potential. In a study of seven Sub-Saharan African (SSA) countries, Van Biesebroeck (2005) finds that it is rare for micro- and small firms in these economies to reach medium or large scale, and that, as in countries outside the region, large firms are the most productive. These firms do not tend to grow and employ more workers over time (La Porta and Shleifer 2014). Thus in contrast to the case of more developed economies such as the United States, firms in SSA which start out small are unlikely to contribute substantially to long run productivity growth.

The predominance of OOMEs—or the lack of larger, more formal firms—also adversely affects wages and job creation. Although in many countries it is believed that smaller firms create more jobs, in some countries they are also disproportionately responsible for even more job destruction. (For the case in the United States, see Neumark, Wall, and Zhang 2009) and Li and Rama (2015).[5] Moreover, a significant body of literature finds that laborers in informal enterprises earn significantly lower wages than those employed in formal enterprises, after controlling for worker characteristics (Montenegro and Patrinos 2014). In Madagascar, only 30 percent of OOMEs employ workers besides their owners, and only 8 percent employ paid laborers. On average, they employ 1.4 workers: 1.0 is the owner, 0.3 are unpaid workers, and 0.1 are paid (authors' calculations using ENEMPSI 2012). Given the observed patterns, there are serious doubts regarding the potential of OOMEs to alleviate poverty and improve productivity.

Nonetheless, a priori it is not clear what the economic losses are, what blocks OOME growth, and what the policy levers are for poverty reduction in countries where they predominate. Microenterprises may be an efficient outcome for some activities and markets. The products and services they produce may differ from those produced by larger firms in ways that are more in line with market demand.[6] Larger firms selling close substitutes may serve a segment of the market, while OOMEs serve another. Entrepreneurs may be more productive given their preferences and the incentive advantages of working for themselves. In larger organizations, the costs of monitoring workers may be high. For a variety of reasons, a dispersion of firm sizes would be expected, and some microenterprises may be efficient in some contexts. Yet if any of these circumstances make OOMEs an optimal outcome, then they should show similar productivity as larger firms. The reasons for the persistent prevalence in poor countries of informal microenterprises (which do not grow into more productive small, medium, or large firms) are not fully understood, and they may be country specific. Grimm, Kruger, and Lay (2011), for example, find that returns to capital at low levels of operation are high in a sample of seven West African countries (with some exceptions), and therefore microentrepreneurs could use internally generated resources to grow, as McKenzie and Woodruff (2006) point out. A full explanation must therefore include impediments to the entry of larger, more productive firms as well as to the growth of micro and small ones. If they were more efficient, large firms would typically drive a significant share of their smaller, less-efficient counterparts out of business while also absorbing the labor released as those firms exit. Barriers to this process can be related to many factors, which require a country-specific diagnosis in order to assess the contribution of causes, such as a lack of infrastructure, an unfavorable investment climate, trade barriers, or difficulties in accessing key inputs such as financial and human capital. Similarly, the constraints that are most binding for OOME growth may be country specific.[7] To shed light on these questions in the case of Madagascar, we therefore focus on the returns to the microentrepreneur, to his or her decision to hire (more) workers, to invest in the enterprise, and to apply his or her own time and effort to the enterprise.

Data and Characteristics of Madagascar's OOME Sector

We utilize data from the *Enquête nationale sur l'emploi et le secteur informel* (ENEMPSI) that was collected in 2012 in two phases, consisting in the first phase of an LFS, and in the second phase a survey of informal enterprises identified through the LFS. For the LFS, a stratified random sample was drawn of over 11,000 households from which 41,000 individuals ages five and over were surveyed. From this sample, a listing of individuals reporting that they owned and operated an informal enterprise was produced, and a representative sample was drawn of such enterprises for the second phase. For the purposes of the survey, informality was defined as not having a statistical number or, in the case of people working "on their own account," as not keeping financial accounts (INSTAT 2013). Only urban-based enterprises were included, whether from large or secondary cities of the country.

The resulting sample was of 5,692 enterprises, of which over 3,968 are owned and operated by a single person with no other workers (single-worker/owner); and 1,724 of which had additional employees, whether unpaid, paid, or both (multi-worker). The activities performed by these enterprises can be classified into the following broad sectors: the primary sector (primarily forestry and forest-related products); industry (manufacturing, construction, mining and quarrying); trade and commerce (wholesale and retail); and all other services (transportation, hotels and food service, household services, information and communication, real estate, professional services, and others). Table 5.2 shows the distribution by gender of single- versus multi-worker OOMEs. Table 5.3 provides breakdown of the sample by type and sector.

The activities represented in the sample do not appear at first glance to be those often efficiently done by individuals working on their own with little capital. Table 5.4 shows the most frequent specific activities performed, by gender and OOME type. As shown, there is significant overlap in activities performed by both single- and multi-worker firms, with both engaged in various types of retail, household services, and mining. In addition to the top 11 most frequent activities shown, OOMEs engage in activities ranging from the sale of motor vehicles to wireless telecommunications, veterinary services, footwear manufacturing, real estate, photography, and automobile rental.

TABLE 5.2: Gender-type Distribution of OOME's

Of which:	Male	Female
Single worker	63.6	76.0
Multi-worker	36.4	24.0
All	44.9	55.1

Source: ENEMPSI 2012.

There is a significant gender dimension to the types of activities pursued. Female-owned businesses are more likely to be engaged in retail, spinning, and textile activities, and male-owned businesses are more likely to be engaged in construction, transport, lumbering, and metallurgy. Although the gender of the entrepreneur does not change the probability of being employed in industry, being male increases the probability of heading an OOME in the primary sector by 86 percent relative to the probability for females, and the other services sector by 12 percent, but decreases the probability of being in the trade sector by 26 percent.[8]

In Madagascar, the majority of OOMEs operate at a very small scale. The level of capital invested is low, but rarely zero: only 1.49 percent of informal firms claim to have zero assets—1.6 percent of single-worker OOMEs and 1.1 percent of multi-worker OOMEs. Mean capital invested in each firm was approximately 736,000 ariary, or US$335.[9] The means and median levels of assets are as shown in table 5.5. Assets are lower for single-worker OOMEs on average, as well as for female-owned ones, but there is overlap in the distributions.

In addition to low levels of capital, the level of employment generated per OOME is very low. Over 70 percent of OOMEs employ only their owners and no other workers, paid or unpaid. Almost all (99.5 percent) employ five or fewer employees, and these proportions are very similar to what they were when those firms started their businesses: 99.5 percent had fewer than five employees, and 76.0 percent claim to have started with

TABLE 5.3: Sample of Informal Microenterprises ENEMPSI 2012 (Not Population Weighted)

	Single-worker/owner		Multiple worker		
Sectors represented	Registered*	Not registered	Registered	Not registered	Total
All	353	3,615	382	1,342	5,692
Industry	26	1,444	58	648	2,176
Primary sector	3	106	7	62	178
Services (except trade)	144	792	91	197	1,224
Trade/commerce	179	1,261	226	438	2,104

*The enterprise counts as registered if it is registered with the commerce department, has a license, has a *carte professionnelle,* or is registered with the social security administration.

TABLE 5.4: Most Frequent Activities by OOME Type and Gender (Percent)

Male-owned single worker	Female-owned single worker	Male-owned single worker	Female-owned multi-worker
Building construction (10.7)	Spinning, weaving, finishing textiles (31.9)	Mining of uranium and thorium ores (15.1)	Retail sale in nonspecialized stores (23.8)
Other land transport (7.9)	Retail sales via stalls and markets (11.1)	Building construction (12.1)	Spinning, weaving, and finishing textiles (12.9)
Mining of uranium and thorium ores (7.8)	Retail sale of food, beverages, tobacco (9.3)	Retail sale in nonspecialized stores (8.5)	Restaurants and mobile food services (10.7)
Retail sale of food, beverages and tobacco (6.8)	Retail sale in nonspecialized stores (8.6)	Lumbering (6.2)	Retail sale of food, beverages, and tobacco (8.5)
Retail sale via stalls and markets (6.8)	Household services (7.2)	Retail sale via stalls and markets (5.9)	Retail sale via stalls and markets (7.5)
Retail trade not in stores, stalls, or markets (3.5)	Manufacture of food products (4.1)	Other land transport (5.4)	Manufacture of other food products (6.4)
Manufacturing of nonmetallic mineral products (3.0)	Manufacturing except fur apparel (3.5)	Retail sale of food, beverages, and tobacco (5.3)	Mining of uranium and thorium ores (4.4)
Supporting transport activities (2.8)	Retail of other goods in specialized stores (3.1)	Retail sale of other goods in specialized shops (4.1)	Manufacturing except fur apparel (2.9)
Other personal services (2.6)	Restaurants and mobile food services (2.9)	Manufacture of nonmetallic mineral products (3.9)	Household services (2.9)
Retail sale of other goods in specialized shops (2.6)	Other personal services (2.6)	Manufacture of other fabricated metal products (2.6)	Retail sale of other goods in specialized shops (2.1)
Wholesale of agricultural raw materials (2.5)	Mining of uranium and thorium ores (2.5)	Extraction of sand stones and clay (2.4)	Wholesale of agricultural raw materials (1.6)

Source: ENEMPSI 2012.

TABLE 5.5: Asset Values by Single- and Multi-worker OOMEs and Gender of Owner (1,000 Ariary)

Gender of owner	Mean			Median		
	Male	Female	All	Male	Female	All
Single worker	608	481	532	50	43	46
Multi-worker	1,344	1,027	1,200	103	113	109

just the one owner/worker. Although 8.0 percent of businesses were created in the past year, most (53.6 percent) were created at least five years earlier, and 34.3 percent at least 10 years earlier. Some OOMEs employ unpaid workers, who are typically family members. The level of employment of unpaid and workers averaged across OOMEs was 35 hours and 24 hours per month, respectively. Multi-worker firms utilized on average 118 hours per month of unpaid labor and 81 hours per month of paid labor. Yet average hours per paid worker were close to or exceeded full time (calculated as 40 hours per week, or 172 hours per month): Paid workers worked on average 1660 per month, and unpaid workers only 22.6 hours per month.

The Impact of Scale: Estimation Method and Results

To understand the economic effects of the OOMEs' scale of operation, we estimate the relationship between firms' monthly cash profits, denoted π, and their use of inputs.[10] We estimate these relationships for all OOMEs, as well as separately for multi-worker firms and single-worker firms, given that single- and multi-worker firms may be fundamentally different in some way. Indexing firms by i, we estimate the following equation:

$$\pi_i = \alpha + \beta f(k_i) + \gamma(g(h_i)) + \theta(X_i) + \xi\hat{\eta}_{iv} + \varepsilon_i, \qquad (1)$$

Where *f(k)* is a polynomial function of *k*, capital invested (or assets), and *g(h)* represents a function of labor hours used. We distinguish between hours of work by the owner-operator, or "head," by unpaid workers, and by paid workers, which combined are represented by the vector h_i. We estimate the functional forms for *f(k)* and *g(h)* by testing the significance levels of higher order terms and dropping those which are not statistically significantly different from zero.[11] This approach has the advantage of allowing some flexibility of functional form, including nonconstant elasticities of profits with respect to inputs, and in contrast to a log specification, of allowing some variables to be less than or equal to zero. We also include X_i, a set of conditioning variables which could affect profitability, including indicator variables for the region of the country, size of urban location (major urban versus secondary urban), and the gender of the owner-operator. We estimate the relationships with this parsimonious set of X_i variables, and with a more comprehensive set of firm characteristics, which includes the years of schooling of the head of the enterprise, the reason the firm gives for pursuing the particular activity, its broad sector (industry, trade, services), the age of the firm, and the age of the owner-operator. In addition, since equation (1) is not a production function, but rather captures the combined effects of scale through both average costs and revenues or prices charged on profits, we attempt to control for demand-side and pricing issues by also including the main type of competitor (large versus small, commercial versus noncommercial).[12]

The term $\xi\hat{\eta}_{iv}$ represents an estimate of unobserved determinants of individual *i*'s decision to own and operate an OOME, which may be correlated with included regressors and π_i. The remaining errors term ε_i is independent and identically distributed. If laborers in the economy sort into sectors, firm types, and types of employment by educational attainment and ability, more (or less) capable workers may sort into OOMEs rather than other means of employment. Alternatively, less capable workers may sort into owning OOMEs because they have more difficulty retaining salaried employment. In either case, unobserved heterogeneity in ability, opportunities, or other characteristics of the entrepreneur may be correlated with thus biasing the coefficient on the owners' time, assets, or other variables. We attempt to address this (selection bias) issue through a two-step estimation strategy which incorporates a multinomial choice model in the first step. In particular, we use the LFS to estimate the probability of an individual *j*'s choice, *v*, of (i) owning and operating a single-worker OOME, (*ii*) a multi-worker OOME, or no OOME, using multinomial probit estimation as shown:

$$pr(y_j = \nu) = \Phi(z_{j\nu}\Gamma). \qquad (2)$$

$\Phi(\ )$ represents the standard normal cumulative distribution function and the error term, η, is distributed according to the standard normal. The vector *z* includes a set of individual and spatial characteristics: the individual's level of education, gender, region, size of city (large or secondary), and a set of family history variables: the years of education of the mother, years of education of the father, the father's job status, and the father's previous sector of employment. In all cases, there is a subset $\tilde{z}_j, \in z_j$ among the family history variables which are excluded from the main equation (1). We then estimate equation (1) using an approximation of η, the generalized residual from equation (2) following Das, Newey, and Vella (2003). In practice, we use a polynomial in the predicted probabilities of each outcome, and we use cross validation techniques to select the preferred specification.[13] In cases where there is no appreciable selection bias in comparison with the ordinary least squares (OLS) estimates, and where the polynomial approximation terms were not jointly significant, we prefer OLS for efficiency reasons.[14]

The results from equation (2), which provides an estimate of the main factors affecting the probability of owning an OOME are summarized in table 5.6. We find that having a father who was a salaried manual laborer or who worked in a nonagricultural sector is predictive of owning some type of OOME; that having more education makes owning such an enterprise less likely and being in a large urban center more likely. Moreover, being male is associated with owning a multi-worker OOME, whereas being female is associated with owning a single-worker OOME.

The estimates of equation (1), shown in table 5.7, provide the following key insights into the issue of (profitability) returns to scale. First, the structure of returns differs substantially between single-worker and multi-worker OOMEs.[15] The returns to both capital and the head's labor are significantly higher for multiple-worker firms, even when firm characteristics such as sector, educational attainment, location, and unobservable factors are taken into account. Moreover, returns to capital are low and diminishing for single-worker OOMEs, on an

TABLE 5.6: Significant Influences on Likelihood of Owning and Operating an OOME

	Single worker OOME	Multi-worker OOME
Father salaried manual laborer	0.447*	0.253
	(2.34)	(1.07)
Father worked in agriculture	–0.385***	–0.279***
	(–6.30)	(–3.68)
Father worked in trade	0.124	0.269**
	(1.78)	(3.19)
Male "= 1 male, = 0, female"	–0.157***	0.192***
	(–4.97)	(4.97)
Age	0.0180***	0.0208***
	(16.53)	(15.81)
Years of education	–0.0471***	–0.0336***
	(–8.76)	(–5.16)
Large urban	0.153***	0.188***
	(4.26)	(4.27)
N	20,065	

*p-value = .10; **p-value = .05; ***p-value=.01; t-statistics in parentheses.
Note: Includes regional indicators and other variables. See annex table 5A.1 for all details.

TABLE 5.7: Estimated Monthly Returns to Firm Inputs

	No firm characteristics		With firm characteristics		
	Single worker	Multiple worker	Single worker	Multiple worker	All
Assets	0.0130***	0.0330***	0.0104**	0.0384***	0.0191***
Assets^2	$-1.37xe^{-10}$**	~	–.000000107*	~	.000000134***
head's hours	1.026***	1.868**	0.743***	1.893**	1.057***
head's hours^2	–0.00120*	–0.00355*	~	–0.00363*	–0.00169**
Unpaid hours		0.985***		0.325**	0.356***
Unpaid hours^2		–0.000848*		~	~
Paid hours		0.494***		0.376***	0.490***
Paid hours^2		–.000266***		–0.000289***	–0.000275***
Male = 1	58.64***	~	83.07***	~	63.72***
Years schooling			11.49***	~	10.69***
Returns to additional year of head's schooling			6.84%	~	4.90%
N	2,332	1,210	2,775	1,355	4,125

* p-value = .10; ** p-value = .05; *** p-value=.01. ~ = not statistically significant.
Note: Includes regional indicators and other variables. Details are reported in annex table 5A.2.

annualized basis averaging 4.5 percent (controlling for firm characteristics), but not for multi-worker OOMEs. Similarly, in the pooled sample, returns to capital are increasing, as shown by the positive coefficient on assets squared, in part due to the increased returns associated with moving from a single- to a multi-worker firm.[16]

These estimates indicate that the returns to capital for single-worker OOMEs fall well below the opportunity cost of capital in the economy. First, they are lower than returns in slightly larger OOMEs, which have an annualized return above 36 percent. The average (bank) lending interest rate at the time was 18 percent, and interest rates on microcredit were 36 percent. Thus, there is a substantial economic loss associated with the market structures observed, where 70 percent of OOMEs are the single-worker type.

Perhaps surprisingly, having more education does not dramatically improve profitability of OOMEs. This may explain why education has a negative selection effect on entry of these entrepreneurs. Returns to schooling conditioning on owning an OOME and on other included variables are positive and significant for single-worker OOMEs, but are not statistically significant for multi-worker ones. Moreover, these returns are low relative to international benchmarks of approximately 10 percent (see Montenegro and Patrinos 2014). In the broader LFS, the estimated returns range from 10.4 to 11.7 percent, depending on the set of conditioning variables, a level close to international benchmarks.[17] However, given that more education is predictive of owning a larger OOME, conditioning on size and other characteristics may understate the actual returns to education.

In addition, in the estimations with firm characteristics, we find that the type of competition faced from the firm's main competitors has an impact on the firm's profitability. Of the four types of competition firms could report—large and small commercial operators and large and small noncommercial operators—competing primarily with a large commercial operator increased profits significantly (see table 5A.3). Smaller entities appear to provide more direct or intense competition with each other than do large commercial entities. Large companies would require a higher return on capital in line with the cost of capital in the economy; they must also pay the costs of formality. Finally, because they enjoy greater market power, their presence may permit OOMEs in the same sector to set higher prices as well.

The structure of returns also varies, as one would expect, by sector. Estimates for industry and trading (sectors with sufficient observations to estimate equation (1) separately) are shown in table 5.8. In industry, returns to capital and the owners' time are again higher for multi-worker firms, although these differences are less pronounced in trading. Male entrepreneurs and those with more education have higher profits in some subsets of the sample, and returns to capital are increasing for industry and mildly decreasing in trading.

Labor Market Frictions

The estimates reported imply that firms not only employ too little capital, but also too little labor, conditioning on their level of assets. First, expanding from a single-worker firm to a multi-worker firm improves the returns to the owner's time and capital appreciably. Second, conditioning on assets, although the returns to paid labor are diminishing, they are still positive for most firms.[18] The estimated profit-maximizing level of paid worker hours per month would be 760 hours, which translates to 4.4 full time-equivalent workers. This is substantially higher than

TABLE 5.8: Estimates of Returns to Productive Factors by Broad Sector

	Industry		Trading	
	Single worker	Multiple worker	Single worker	Multiple worker
Assets	.0199***	.261**	.105~	0.122 **
Assets^2	~	–.0000826**	–0.0000425~	–0.0000183 **
Assets^3	~	5×10^{-9}**		~
Head hours	.56***	1.041**	0.340~	.58~
Head hours^2		~		
Unpaid hours		~		.475 **
Unpaid hours^2		~		
Paid hours		~		–.09~
Paid hours squared		~		
Male = 1	79.8***	~	103*	290.7***
Schooling of head	8.431***	~	~	32.9***
N	1,269	285	731	384

TABLE 5.9: Marginal Profit of Paid Labor and Wages Paid

	Percentile of observations		
	25th	50th	75th
Marginal (profit) product	.263	.357	.414
Survey estimates average wage	.213	.400	.714
Friction/wage	1.23	.89	.58

current full-time employment of on average 81 paid hours and .47 paid workers for multi-worker firms.

The gap between actual and optimal labor employment suggests an important labor market friction, typically due to the costs of finding, monitoring, managing, and dismissing workers (Rogerson, Shimer, and Wright 2005). The relationship between the unexploited marginal profit and wages suggests that the friction diminishes as a percentage of average wage as wages (and labor productivity) increase (see table 5.9). To assess whether such costs vary with firm characteristics, we examine correlations of these with the number of hours of paid labor used. All else equal, male-owned firms hire 330 more paid labor hours per month. Firms headed by an individual whose father was a "boss" hire 250 more labor hours, whereas those whose father was employed in agriculture hire 120 hours less. Being located in a large urban area is associated with hiring 146 hours more of labor. Finally, for each year of completed education, heads of firms hired 29 more hours of paid labor per month. While the difference between single-worker OOMEs, which have 4.5 years of education on average, and multi-worker OOMEs, at 4.8 years, is not very large, informal firms with paid workers have an average of 6.0 years of schooling. Based on these correlations, it appears to be more costly for female owners to hire paid workers, and having more education and training appears to reduce the costs of hiring paid workers. See annex table 5A.6 for more details.[19]

Profit Elasticities and Low-Productivity Traps

To assess whether there are increasing (profit) returns to scale, we compute the elasticities of profits with respect to each factor of production and for all factors added together. If these elasticities are positive, then firms could increase their profits simply by scaling up, using the available technologies.[20] As shown in table 5.10, with the exception of paid worker hours and owner hours at or above the 90th percentile of the factor's use, all elasticities are positive. Moreover, the profit elasticity with respect to capital assets is generally increasing as one moves up the distribution. A 1 percent increase in the firm's asset base (or capital) will contribute only slightly to profits at the low end of the scale, but as the firm's capital expands, it increasingly adds to profitability, holding other factors costs constant. However, because these gains are not offset by the cost of borrowing, single-worker firms would have to achieve a certain scale to make borrowing for investment purposes worthwhile. With respect to labor inputs, to maximize profits would entail hiring paid labor at approximately the 75th percentile level for multi-worker firms, and above the 90th percentile of all OOMEs.

These positive elasticities, combined with the coefficients on returns reported in table 5.7, provide evidence of a major allocative inefficiency and a low-productivity "trap." Too much capital and labor are employed in single-worker OOMEs, which have low returns to both factors, and among multi-worker OOMEs there remain unexploited economies of scale. These results contrast with those found in several West African countries. Grimm, Kruger, and Lay (2011) find that returns to capital at low levels of operation are as high as 70 percent per month in all urban areas but one (Lomé) and fall with the level of capital. Therefore, their results do not support a hypothesis of a low-productivity trap: for this to occur, the returns to capital at a very small scale would be too low to enable poor microentrepreneurs to grow their businesses through internally generated resources. Yet in Madagascar, where returns are lower at the lowest levels of investment and generally rising with scale, this hypothesis appears to have validity.

The Returns to Microentrepreneurs' Labor

Many OOME heads responded to the survey by saying that they chose to operate a business to improve their incomes. (See the "Introduction" to this chapter.) To see what the effects on labor income of this mode of (self-) employment are, we compare average labor income per month among different categories of workers as captured

TABLE 5.10: Elasticity of Profits with Respect to Factors of Production by Percentile of Inputs Used (Unweighted)

	10th	25th	50th	75th	90th	95th
Assets						
All	0.0005	0.0025	0.0070	0.0307	0.1347	0.4635
Single worker	0.0002	0.0012	0.0032	0.0132	0.0681	0.2119
Multi-worker	0.0026	0.0074	0.0153	0.0705	0.3778	0.7901
Paid worker hours						
All	0.1007	0.2414	0.1797	0.1667	0.1492	–0.2277
Multi-worker	0.0744	0.1713	0.1183	0.0808	–0.0500	–0.4821
Unpaid worker hours						
All	0.0710	0.1248	0.2133	0.3002	0.2344	0.3875
Multi-worker	0.0648	0.1139	0.1947	0.2740	0.2139	0.3536
Owner hours						
All	0.3687	0.4154	0.3839	0.2124	0.0787	–0.0722
Single-worker	0.4294	0.5122	0.6752	0.6815	0.8842	0.7680
Multi-worker	0.4020	0.3972	0.3413	0.0744	–0.2803	–0.4699
All factors, total and *using no negative factors*						
All	0.5409	0.7841	0.7839	0.7100	0.5971	0.5510
All (no neg)	*0.5409*	*0.7841*	*0.7839*	*0.7100*	*0.5971*	*0.8510*
Single-worker	0.4297	0.5133	0.6784	0.6947	0.9523	0.9799
Multi-worker	0.5438	0.6899	0.6696	0.4996	0.2613	0.1917
M-w (no neg)	0.5438	0.6899	0.6696	0.4996	0.5917	1.1437

Note: Elasticity estimates are computed using the coefficients estimated and reported above and locally smoothed values of the independent variables and associated profits in the sample within 5 percent on either size of the indicated percentiles.

in the LFS. Because wages are not observed for owners of OOMEs, we estimated a shadow wage equal to the average estimated contribution to profits of the OOME heads' labor using the coefficients estimated and the level of OOME labor applied for each firm. We then estimate a linear wage equation (3) in two steps, as follows:

$$w_j = \alpha + \beta X_j + \mu \hat{\eta}_j + \varepsilon_j, \qquad (3)$$

where w_j represents the monthly labor income of individuals in the LFS, indexed by j. X_j includes the age, level of education and training, region, size of urban center, sector of employment, and an array of family history variables for individual j. $\hat{\eta}_j$ is a polynomial approximation to the generalized residual from equation (2), which in this case captures unobserved heterogeneity affecting labor income—in the case of OOME heads, unobservable ability or opportunity. We include binary indicator ("dummy") variables for those owning and operating an OOME, for workers in the public administration, workers in agriculture, industry, trade, and services, and for paid employees (not heads) of OOMEs receiving a positive wage. The excluded category of workers are therefore those workers most like the OOME head, but who work for a wage in the private sector. The results are summarized in table 5.11. All else equal, heads of single-worker OOMEs are paid significantly less than their counterparts who work in the private sector. The loss in income of 60,000 ariary per month is substantial relative to the mean earnings of urban-based workers of 101,000 ariary (INSTAT 2013). When one estimates these relationships using OLS, however, the effect of single-OOME ownership is not significant, implying that their unobserved ability increases their actual earnings relative to comparator workers. In contrast, heads of multi-worker OOMEs make on average 68,000 ariary per month more than their counterparts in regular employment. Their business assets and scale appear complementary to their labor effort, even when one conditions on family history and unobservables. Although heading a single-worker OOME represents a major inefficiency, operating a larger multi-worker OOME can be a gainful choice for those willing and able to do so.

TABLE 5.11: Wage/Average Returns to Labor (1000 Ariary per Month)

	Bias corrected	OLS
Single-worker OOME	–59.57**	–3.90
	(–2.85)	(–1.06)
Multi-worker OOME	68.47***	53.43***
	(10.86)	(11.58)
Employee of OOME	–12.00***	–11.80***
	(–3.40)	(–3.51)
Received vocational training	76.56***	76.53***
	(8.41)	(14.62)
Male	31.59***	29.87***
	(13.83)	(12.49)
Years of education	9.042***	8.819***
	(17.53)	(21.94)

Note: t-stats (bootstrapped as appropriate) in parentheses

Toward a Unified Theory of OOMEs: Asymmetric Information and Incomplete Markets

The pattern identified in returns to capital and labor estimated from the sample of Malagasy OOMEs presents a puzzle. That is, in the presence of increasing returns, even if firms start small, over time they could invest a small amount more in each period, employing more labor as assets expand, and thus achieve higher profitability over time.[21] This is true even if entrepreneurs are risk averse.[22] Those that achieved greater scale could employ more workers and drive the less productive firms out of business, while offering higher wages as well. Thus, even if many entrepreneurs face borrowing constraints, if the most able can access financing and there were otherwise relatively free entry and exit, resources (capital and labor) would flow to the most productive enterprises able to grow more quickly. Credit constraints can constitute a barrier to entry if these constraints apply to almost all entrepreneurs. However, they do not provide a complete explanation for the inefficient market structure observed in Madagascar. With economies of scale, saving small amounts of retained earnings each period would result in firm and productivity growth. Moreover, even microenterprises that cannot borrow could combine their capital to create larger firms and achieve greater economies of scale. Thus, the apparent persistence of low-productivity OOMEs requires a more multifaceted explanation of the failure of OOMEs to grow despite the possibility of financing growth through credit, internal resources, or mergers and partnerships.

We propose a hypothesis that relies upon a combination of market failures to produce a persistent ("steady") state of low productivity for OOMEs. First, there must be a constraint to the competitive pressures that OOMEs face from larger, more competitive firms, both in serving the market and in retaining labor. Second, OOME owners must be poor—that is, they must have low levels of wealth relative to the range where returns are highly increasing. Third, financial markets must fail to serve a significant portion of entrepreneurs. Finally, and relatedly, due to information asymmetries, it must be costly for entrepreneurs to merge their resources through partnerships (or other similar financial or ownership arrangements) to take advantage of scale economies.

To illustrate in a dynamic setting how these circumstances could produce a low-productivity trap, we consider a simple model of the entrepreneur's decision of how much to invest and how many entrepreneurs with which to partner. We assume that competitive pressures from larger entities are limited for reasons outside our model.

Entrepreneurs have available wealth in period t denoted x_t. The value of this wealth can be captured in a value function V, as follows:

$$V(x_t) \equiv E_t \sum_{t=1}^{T} \mu(c_t) = \mu(c_t) + E_t V(x_{t+1}), \qquad (4)$$

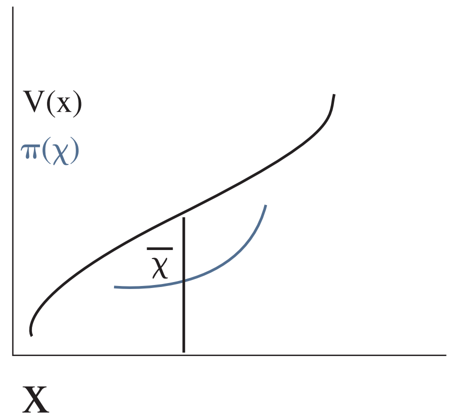
where $\mu(c_t)$ denotes contemporaneous utility. In each period t there are two decisions: first the entrepreneur decides the number of owners, B, to have (re-)invest in the enterprise, knowing that each will invest an optimal level for them, k_t^* in the second stage. Solving the problem backward, each entrepreneur must decide how much of available resources to consume (c_t), how much to save in liquid assets (a_t), and how much to invest in the enterprise, (k_t), which can include paid labor, raw materials, or capital assets. To simplify the model, assume that all owners invest and intend to split profits evenly, have equal wealth, face the same constraints, and have the same utility functions. Total capital invested in t will thus equal $K_t \equiv k_t B_t$.

Because there is no credit market, in all periods, entrepreneurs must have nonnegative assets for all t: i.e., $x_t - k_t - c_t = a_t \geq 0$. If the entrepreneur is borrowing-constrained in period t in the sense that he would be better off borrowing if he could in period t, then his household consumes and invests all resources, that is, $c_t = x_t - k_t$, and he sets savings (a_t) to zero.[23] In this case, in period t the entrepreneur sets the marginal utility of consumption equal to the expected value in the subsequent period of resources earned through the enterprise (in t+1). These resources are thus determined by the marginal value product of k_t. In particular, for all t, and optimum number of enterprise owners B^* and market demand d, the entrepreneur sets consumption c_t such that

$$\mu'(c_t) = E_t\left[\frac{dV_{t+1}(\tilde{\pi}_{t+1})}{dx_{t+1}}\frac{d\tilde{\pi}(k_t B^*, d)}{dk_t}\varepsilon_{t+1}\right], \qquad (5)$$

where ε_{t+1} is a random shock to profits centered at 1. We assume the standard properties of the utility function $-\mu'(c) > 0$, and $\mu''(c) < 0$. That is, the marginal utility of consumption is declining with the level of consumption and is high when consumption is low, as is typically assumed. Under standard assumptions, with concave profit (that is, $\pi''(K) < 0$), $V(x_t)$) is also concave. In this case, however, this is not guaranteed. Profits, $\tilde{\pi}$, are an increasing function of total investments made, $k_t B^*$. In these circumstances, there is a range of x for which $E[V(x_{t+1})]$ is concave, and a range for which it *may* be convex.[24] More formally, if there is a range of total firm capital, kB for which $\pi'' > 0$, then it is possible for $\phi(x_t) \equiv \frac{dV_{t+1}}{dx_{t+1}}\frac{d\pi(k_t B^*)}{dk_t}$ to be increasing in x_t—i.e., convex. Under certain functional forms, including if $V''' = \pi''' = 0$, there is a value of x, denoted $\bar{x}$, such that $\phi' < 0$ for $x < \bar{x}$, and $\phi' > 0$ for $x > \bar{x}$. This implies that the expected value function is an upside-down S-shape, as shown in figure 5.3. The implication of the concavity of $V(x)$ for $x < \bar{x}$ despite increasing returns to K is that since the marginal utility of c is high, entrepreneurs choose a level of K_t such that there are unexploited increasing returns to scale (in expectation) and for which $\pi'(K)$ is low. How probable it is that the entrepreneur adopts a low-productivity strategy, with low returns to capital, therefore, depends on his level of wealth, x, relative to $\bar{x}$.[25] The poorer the entrepreneur is, the more likely she is to set k below the range where returns are high (See figure 5.3).

FIGURE 5.3: Shape of Profit and Value Functions



Solving now for the first stage of the entrepreneur's problem, taking as given that any potential partners will set $k_t = k^*(x_t)$ in the second stage, the entrepreneur chooses B, according to the following optimization problem to maximize his individual benefit $\tilde{\pi}$:

$$\max_B \tilde{\pi}(k^*, B) = \left[\frac{\pi(k^*B; d)}{B} - m(B-1)\right]$$
$$\text{for } B \geq 1.$$

There is a monitoring cost of m, which increases linearly with the number of owners, incurred to ensure that other owners do not withdraw excessive resources from the firm. If the expression in brackets is concave, the solution to the first order condition gives optimal B as:

$$B^*(x,m,d) = \frac{\pi'k^* + \sqrt[2]{(\pi'k^*)^2 - 4m\pi}}{2m}.$$

For a solution with $B > 1$ to exist, $(\pi'k^*)^2 - 4m\pi \geq 0$ —that is, the incremental profits from an additional investor must be high relative to current profits and monitoring costs. Otherwise, $B^* = 1$.

Differentiating the first order condition further, one obtains that $\frac{dB^*}{dm} < 0$. As expected, the optimal number of partners is lower the higher is the cost of monitoring them. Moreover, $\frac{dB^*}{dk^*} > 0$ if $\frac{d^2\pi}{dk^{*2}} > 0$ in the relevant range. That is, the number of partners is increasing in the level of capital all partners are willing to invest. Monitoring carries a fixed cost, and if potential entrepreneurs are too poor, they tend to work alone. Each entrepreneur knows that other possible partners are, like himself, drawing their household consumption resources from

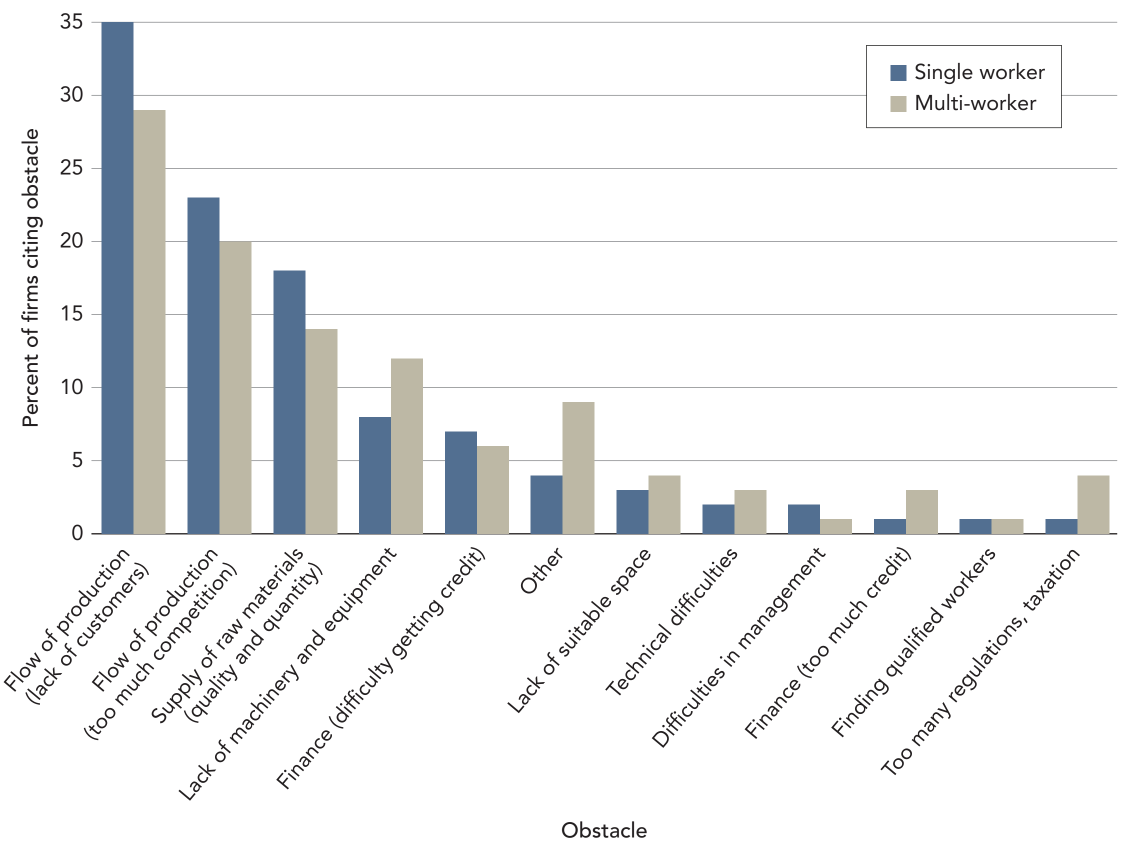
TABLE 5.12: Client Types and Firm Profits

Main Clients	Share of firms	Average value added by main client type
Public and semi-public sector	0.08	322.6
Large private company (commercial)	1.38	305.7
Small business (noncommercial)	3.78	269.5
Large private company (noncommercial)	1.08	220.3
Small business (commercial)	16.52	182.3
Households	77.01	164.8

the revenues of the firm, and that given entrepreneurs' poverty, there is a high marginal utility of consumption. This may only increase the costs of monitoring other owners' actions and firm resources. Therefore, partnerships are unlikely to form. Among informal enterprises in Madagascar, as is generally observed elsewhere, shared ownership is rare: less than 1 percent of OOMEs are limited liability companies. For similar reasons, offering credit—which from the lender's perspective looks similar to becoming a partner as it requires the assumption of risk and monitoring costs—will typically not be worthwhile.

Finally, there is an important market-level effect of the general lack of demand, which could be introduced in a general equilibrium extension of this (partial equilibrium) model. In a situation where incomes are low (that is, almost all businesses are too small), and entrepreneurial wealth and labor earnings are low, equilibrium market demand for goods and services is reduced, as are opportunities to offer more differentiated goods and services (Greenwald and Stiglitz 1990). As shown in table 5.12, for 77 percent of OOMEs the main client types are households, followed by other small businesses at 16.5 percent. The type of client matters, as shown in table 5.12. The average value added for OOMEs is highest among those able to serve the public sector and large private companies. Perhaps this explains why, from each entrepreneur's perspective, rather than a lack of credit,

FIGURE 5.4: Most Important Obstacles to OOMEs' Business



OOMEs claim that their main constraints are a lack of demand. As shown in figure 5.4, the most important obstacle to OOMEs that entrepreneurs most frequently cite is the lack of customers, followed by too much competition.

Madagascar's Weak Enforcement and Monitoring Infrastructure

Madagascar's financial markets are relatively undeveloped in part because they lack the supporting information and enforcement infrastructure required to efficiently screen borrowers and enforce repayment. In 2012, credit registries covered over 0.1 percent of the adult population; credit bureaus covered zero percent of the population, and the strength of legal rights to enforce repayment rated only 2 on a 10-point scale in Doing Business's *Getting Credit*. Credit was also costly in 2012, with bank-lending interest rates averaging approximately 18 percent, and real interest rates of 11.6 percent (according to central bank data).[26] At the same time, microcredit had been growing. Launched in 1990, Madagascar's microfinance sector had about 31 players in 2012, which included state, foreign investor, and donor-supported initiatives, operating under a legal framework and regulated by Madagascar's Central Bank. The average lending rate was 36 percent—a rate almost equivalent to the estimated return on capital in the OOME sector. *Given these realities, for single-worker OOMEs, borrowing small amounts would not be profitable*. Moreover, formal creditors would be unlikely to lend to OOMEs without their adopting more transparent accounting standards and/or offering high collateral relative to loan values in order to reduce the monitoring costs *m* relative to loan values.

As a result, use of credit by Madagascar's OOME sector is extremely low. A full 92.5 percent of OOMEs received their assets through a gift, inheritance, or own savings and only 1.2 percent of them through some type of loan. Moreover, only 3.6 percent of single-worker OOMEs and 3.8 percent of multi-worker OOMEs utilized some type of credit to finance the operations of their enterprise. Of all credit transactions reported, most were informal: 48 percent were from family or friends, 20.6 percent from a microcredit institution, 15 percent from suppliers, 9.2 percent from customers, 3.5 percent from money lenders, and 2.0 percent from banks. Moreover, in a regression of enterprise assets on other characteristics, the education of the entrepreneur's mother and father increased his or her enterprise assets, as did being male (conditioning on sector and region). The age of the establishment was not significant, suggesting that firms' assets are typically not accumulated over time.

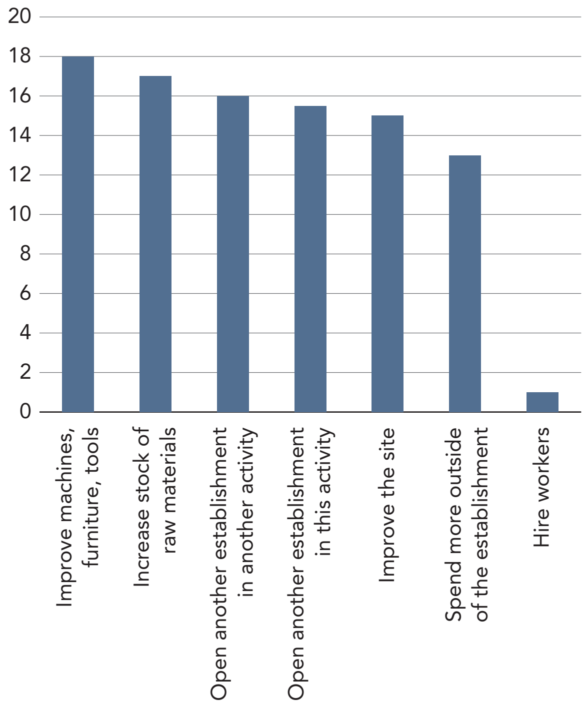
Although OOMEs rate demand-side issues as their most serious obstacles, in the hypothetical situation that they would be granted access to finance, their priority would be to invest in capital. As shown in figure 5.5, the most frequent response entrepreneurs gave to the question "What would be your priority if you could benefit from credit for your activity?" was to improve the quality of machinery and tools, followed by increasing the stock of raw materials and opening another establishment. A very small percentage responded that they would hire more workers. These responses are consistent with the fact that returns to capital are increasing, whereas the estimated marginal product of hiring workers is diminishing and involves important transactions costs.

Formal Registration

The debate regarding the potential of microenterprises is closely intertwined with questions regarding the costs and benefits of firm formality. Thus, a firm characteristic of policy interest is the decision to register one's firm with the authorities. Because the intention in the survey was to sample informal firms, these data provide only a partial perspective on the issue.

Nonetheless, some firms in the survey were more formal than others. Surveyed firms were asked whether they had any of the following types of registration: registration with Ministry of Commerce or the social security fund or possession of a license or professional identification card.[27] If we consider a firm with at least one of these forms of registration, only 9 percent of OOMEs were registered in 2012: 5.7 percent of single-worker OOMEs and 9.3 percent of multi-worker OOMEs. The

FIGURE 5.5: Frequency of Response to "What would be your priority if you could benefit from credit for your activity?"



Source: ENEMPSI 2012

TABLE 5.13: Population of Informal Enterprises by Sector and Type

	Percentage of total	Of which: single worker	Of which: registered
Industry	43.1%	71.8%	1.7%
Primary	4.2%	65.9%	4.1%
Other services	17.4%	74.2%	16.2%
Trade	35.3%	67.9%	15.8%

Source: ENEMPSI 2012.

breakdown by broad sector is shown in table 5.13. Those that did so also achieved a higher level of profitability on average. However, this may not be due to the *effect* of registering: those firms which register may be run by the most capable individuals, with more opportunities for which registration is a requirement, and thus the coefficient on registration from an OLS regression could be biased. To address this issue, we estimate equation (1), this time including a binary indicator variable for whether or not the firm is registered. We first estimate equation (2) with the outcomes v set either to the decision to operate an unregistered OOME of the appropriate type (single- or multi-worker) or a registered OOME of that type and include the variable capturing registration. Again, we use the method of Das, Newey, and Vella (2003) to address possible bias in this coefficient.

In the first stage (equation 2), we find that the sector and type of work of the individual's father, the father's schooling, and the age and gender of the head of the OOME are significant determinants of the decision to register. In particular, having a father who was more educated makes it more likely that the enterprise will be registered, as will having a head who is older or male. (See annex table 5A.6)

The effect of registration is found to be positive and statistically significant for single-worker OOMEs, but is not for multi-worker firms, with other significant results summarized in table 5.14. The bias correction terms are significant, and the coefficients on the variable registered differ significantly when these terms are included (vis-à-vis OLS), suggesting that the effects of registration are attributable to some extent to unobserved heterogeneity in either ability or opportunity.

Whereas the act of registering does not improve profitability on its own, it appears to be associated with access to services and markets that do. Registered firms are more likely to have electricity: 28.3 percent versus only 6.4 percent of unregistered firms. Also, 31 percent have a telephone as opposed to only 6.5 percent of unregistered firms. They are more likely to have a computer (4 percent) or Internet connection (2 percent) than are unregistered firms (less than 1 percent in both cases), although the level of use of these technologies is very

TABLE 5.14: Statistically Significant Firm Characteristics for Profits

Variables	Multi-worker	Single-worker
Registration	~	100.5**
Has electricity	37.3*	40.47**
Fixed locale	~	~
Exporter	245.2**	249.7**

Note: Regional and city-size dummies, other characteristics, and family history variables included.
*p=.10; **p=.05; *** p=.01. ~ = not statistically significantly different from 0. In all cases, cross-validation was used to select specification of bias-corrected estimation.

TABLE 5.15: Percentage of Firms Using Different Types of Locations, Registered versus Unregistered

	Registered	Unregistered
Travelling, improvised on road or public market	10.3	21.7
Stationary on the road	6.5	6.7
Vehicle	14.9	1.2
At home (either fixed or unfixed installation)	14.9	40.2
Fixed locale in public market, workshop, shop/boutique, or restaurant	52.1	25.7

TABLE 5.16: Correlation of Firm Registration and Main Client Type

	Probit estimation coefficient
Public sector clients	0.736*
	(1.73)
Large private commercial client	0.332*
	(1.77)
Assets	3.39e-08***
	(8.67)
Large urban	0.539***
	(10.95)
Direct exporter	0.797*
	(1.76)
N	5338

Note: t statistics in parentheses.

low in general. They are also more likely to have a fixed locale (table 5.15). But perhaps the greatest benefit to registration, however, is improved ability to serve higher value customers. As shown in table 5.16, there is a correlation between registration and serving the highest value clients—public entities, large private companies, and export markets.

Conclusions and Agenda for Further Research

Analysis of the patterns of returns, investment, and employment on the part of owner-operated microenterprises in Madagascar suggests both potential and limits to their ability to contribute to poverty reduction in the country. The current structure of markets, characterized by the prevalence of a large number of unproductive informal microenterprises, results in a substantial misallocation of both capital and labor. The returns to invested capital for the tiniest of these enterprises are below the opportunity cost of capital in the economy, and returns to the entrepreneur's labor are significantly lower than wages of similar individuals in the labor market.

Although owner-operated microenterprises are motivated to invest more to raise their incomes, they are constrained by competition among similar enterprises, by low demand and investor capacity in a context of widespread poverty, difficult access to export markets, and high transactions costs in capital and labor markets.

Although Madagascar's financial markets are undeveloped, the very nature of owner-operated microenterprises increases the importance of high transaction costs and limits the markets' potential to grow. In particular, the high cost of monitoring prospective partners or borrowers relative to profit levels, exacerbated by the lack of formal separation between household and enterprise finances and the risk aversion of poor entrepreneurs, makes this problem difficult to solve using conventional credit markets. Moreover, the costs of microfinance exceed the returns to capital at the smallest scale of operation. Monitoring and control technologies can be expensive, and they are typically more difficult to implement for informal firms precisely because they are informal; these firms lack the practice of keeping clear and precise accounts which can be checked and audited. Because firms with the scale and technologies needed to compete against the market of low-productivity informal firms also face barriers to entry, an equilibrium market configuration persists, with many low-productivity firms competing for the same limited markets.

While there appears some potential to foster the growth of OOMEs that have already achieved a certain scale, to significantly reduce the misallocation of capital and labor misdirected to single-worker OOMEs in particular would require an alleviation of constraints inhibiting investment by larger, more sophisticated firms, which can offer more productive employment and stimulate greater demand and better access export markets. At the same time, given these constraints, the absence of larger, more efficient firms that could stimulate a restructuring of markets and increased labor productivity, incomes, and demand helps to perpetuate the low productivity "trap" indicated by our results. Madagascar's investment climate suffers from political instability, but also from a poor state of its infrastructure, poor access to external markets, and a difficult investment climate. These barriers would need to be lowered for investments in high-potential OOMEs to bear fruit. Moreover, unless the labor market functions efficiently, investments may not bring the jobs required to lift substantial numbers of the population out of poverty. Thus, a greater understanding of the sources of high-transaction costs is needed, so that policies and institutional innovations can be explored to reduce the frictions inherent in informal and formal labor markets.[28]

Annex 5A. Tables

TABLE 5A.1: First Step Estimates: Single- versus Multi-worker OOME

	Single worker OOME	Multi-worker OOME
Father's level is upper, engineer and similar (salaried)	–0.095	0.134
	(–0.40)	(0.48)
Father's level is middle management, foreman (salaried)	0.111	0.192
	(0.55)	(0.78)
Father's level is skilled worker (salaried)	0.192	0.256
	(1.01)	(1.10)
Father semi-skilled worker	0.192	0.140
	(0.97)	(0.57)
Father salaried manual laborer	0.447*	0.253
	(2.34)	(1.07)
Father was unsalaried head	0.308	0.437
	(1.60)	(1.88)
Father independent worker	0.302	0.297
	(1.69)	(1.35)
Father apprentice	–0.708	–0.0170
	(–1.00)	(–0.02)
Father worked in agriculture	–0.385***	–0.279***
	(–6.30)	(–3.68)
Father worked in industry	0.0491	0.139
	(0.72)	(1.66)
Father worked in trade	0.124	0.269**
	(1.78)	(3.19)
Father's years of education	–0.0106	–0.00673
	(–1.58)	(–0.82)
Mother's years of education	–0.00807	–0.00175
	(–1.12)	(–0.20)
Male "= 1 male, = 0, female"	–0.157***	0.192***
	(–4.97)	(4.97)
Age	0.0180***	0.0208***
	(16.53)	(15.81)
Years of education	–0.0471***	–0.0336***
	(–8.76)	(–5.16)
Large urban	0.153***	0.188***
	(4.26)	(4.27)
Includes region dummies (not shown)		
N	20065	

Note: t statistics in parentheses. ***p = .01; **p = .05; *p = .10.

TABLE 5A.2: Value Added Estimation without Firm Characteristics

	Single Worker		Multi-worker	
	2-step	OLS	2-step	OLS
Assets	0.0130***	0.0149***	0.0330***	0.0319***
	(2.87)	(3.77)	(5.30)	(5.85)
Assets squared	–0.000000137**	–0.000000154***	5.74e–08	–1.16e–09
	(–2.23)	(–2.75)	(0.98)	(–0.02)
Hours head	1.030***	1.059***	1.868**	1.315*
	(4.02)	(4.86)	(2.14)	(1.82)
Hours head squared	–0.00120*	–0.00145**	–0.00355*	–0.00208
Squared head monthly working hours	(–1.78)	(–2.52)	(–1.70)	(–1.21)
Paid hours			0.494***	0.552***
			(3.05)	(4.18)
Paid hours squared			–0.000266***	–0.000320***
			(–4.32)	(–6.04)
Unpaid hours			0.985***	0.405*
			(2.78)	(1.91)
Unpaid hours squared			–0.000848*	–0.0000426
			(–1.94)	(–0.25)
= 1 if large urban	91.32***	67.71***	73.69	84.80**
	(5.79)	(5.39)	(1.38)	(2.07)
'1 = male	58.62***	88.19***	11.35	26.21
	(3.77)	(7.11)	(0.21)	(0.67)
Includes region dummies (not shown)				
N	2332	3347	1210	1592

Note: t-statistics in parentheses. *p<0.10; **p<0.05; ***p<0.01.

TABLE 5A.3: Estimates of Determinants of Profits, with Firm Characteristics

	Single worker	Multi-worker	All
Total assets (1000 ariary)	0.0104**	0.0384***	0.0191***
	(2.54)	(14.75)	(5.94)
Total assets squared	–0.000000107*		0.000000134***
	(–1.90)		(4.16)
Boss working hours	0.743***	1.893**	1.057***
	(3.25)	(2.50)	(3.66)
Boss working hours squared	–0.000707	–0.00363**	–0.00169**
	(–1.18)	(–1.97)	(–2.29)
Paid worker hours		0.376***	0.490***
		(2.62)	(6.21)
Paid worker hours squared		–0.000289***	–0.000275***
		(–5.05)	(–7.76)
Unpaid worker hours		0.325**	0.356***
		(2.32)	(4.92)

(continued)

	Single worker	Multi-worker	All
Years of schooling, head	11.49***	7.046	10.69***
	(6.13)	(1.14)	(4.48)
Exporter	264.3**	77.15	189.0
	(2.55)	(0.27)	(1.53)
Competitors small commercial	–26.38	–419.0***	–191.1***
	(–0.56)	(–2.93)	(–3.27)
Competitors large noncommercial	–65.47	–549.9***	–265.2***
	(–1.03)	(–3.04)	(–3.45)
Competitors small noncommercial	–53.39	–386.7***	–200.2***
	(–1.10)	(–2.60)	(–3.32)
Product selected for family tradition	10.20	208.7**	69.40**
	(0.38)	(2.32)	(2.04)
Product selected business know	12.31	115.9	38.86
	(0.52)	(1.42)	(1.28)
Product chosen for profit	83.58***	304.9***	147.4***
	(3.34)	(3.69)	(4.64)
Product chosen for stable revenues	34.13	147.1	69.91*
	(1.03)	(1.50)	(1.74)
Belongs to organization of producers	53.38	369.8***	158.2***
	(1.35)	(3.19)	(3.27)
Industry	73.11*	4.998	46.00
	(1.90)	(0.04)	(0.98)
Service	78.62*	–57.84	45.15
	(1.95)	(–0.46)	(0.90)
Trade	111.8***	113.0	111.9**
	(2.88)	(0.97)	(2.34)
Large urban	33.94**	45.38	32.60*
	(2.37)	(0.96)	(1.79)
Male	83.07***	25.76	63.72***
	(6.09)	(0.59)	(3.75)
Father boss (no salary)	–3.304	178.9**	81.74**
	(–0.10)	(2.12)	(2.18)
Father upper salaried	–3.376	777.0***	307.4***
	(–0.05)	(3.85)	(3.71)
Father worked in agriculture	–33.75**	–92.75*	–57.24***
	(–2.28)	(–1.94)	(–3.05)
Constant	–45.12	595.6**	169.0*
	(–0.57)	(2.26)	(1.68)
N	2775	1355	4125

Note: *t* statistics in parentheses. *** p = .01; **p = .05; *p = .10.

TABLE 5A.4: Multinomial Probit Estimates of Determinants of Registration

Dependent variables	Whole sample: Choice 1: Own, not registered	Whole sample: Choice 2: Own, registered	Single-worker OOME: Choice 1: Own, not registered	Single-worker OOME: Choice 2: Own, registered	Multi-worker OOME: Choice 1: Own, not registered	Multi-worker OOME: Choice 2: Own, registered
Father upper level salaried position	0.0450	–0.263	–0.0173	–0.718	0.116	0.342
	(0.20)	(–0.74)	(–0.07)	(–1.64)	(0.37)	(0.61)
Father was middle manager/ foreman	0.223	–0.186	0.165	–0.432	0.232	0.209
	(1.16)	(–0.57)	(0.77)	(–1.16)	(0.87)	(0.39)
Father skilled salaried	0.257	0.0368	0.231	–0.295	0.171	0.467
	(1.43)	(0.12)	(1.15)	(–0.84)	(0.68)	(0.89)
Father semi-skilled salaried	0.346*	–0.734**	0.322	–0.848**	0.246	–0.449
	(1.84)	(–2.10)	(1.54)	(–2.12)	(0.94)	(–0.77)
Father salaried manual laborer	0.494***	–0.0154	0.539***	–0.396	0.0988	0.411
	(2.72)	(–0.05)	(2.67)	(–1.07)	(0.39)	(0.77)
Father was head (no salary)	0.399**	0.238	0.297	–0.0612	0.352	0.596
	(2.19)	(0.76)	(1.46)	(–0.17)	(1.41)	(1.13)
Father worked on own (unsalaried)	0.358**	0.107	0.337*	–0.187	0.187	0.486
	(2.11)	(0.36)	(1.77)	(–0.56)	(0.80)	(0.95)
Father was unsalaried apprentice	–0.769	0.0595	–0.573	–10.43	–10.17	1.033
	(–1.07)	(0.08)	(–0.79)	(–0.00)	(–0.00)	(1.19)
Father worked in agriculture	–0.358***	–0.367***	–0.350***	–0.351***	–0.157*	–0.251**
	(–6.09)	(–3.87)	(–5.51)	(–2.88)	(–1.82)	(–2.06)
Father worked in industry	0.0941	0.0335	0.0242	0.0640	0.197**	–0.0478
	(1.43)	(0.32)	(0.34)	(0.49)	(2.08)	(–0.34)
Father worked in trading	0.146**	0.289***	0.0822	0.0774	0.170*	0.376***
	(2.16)	(2.84)	(1.13)	(0.58)	(1.73)	(2.97)
Father's years of education	–0.0133**	0.00258	–0.0120*	0.00250	–0.00873	0.00479
	(–2.05)	(0.25)	(–1.70)	(0.19)	(–0.93)	(0.37)
Mother's years of education	–0.00562	–0.00854	–0.00868	–0.00411	0.00374	–0.00876
	(–0.81)	(–0.80)	(–1.15)	(–0.29)	(0.37)	(–0.65)
Male = 1	–0.1000***	0.256***	–0.246***	0.226***	0.224***	0.275***
	(–3.31)	(5.01)	(–7.44)	(3.35)	(5.14)	(4.23)
Age	0.0192***	0.0206***	0.0150***	0.0162***	0.0163***	0.0174***
	(18.44)	(11.86)	(13.29)	(7.12)	(11.13)	(8.01)
Years of education	–0.0622***	0.0298***	–0.0543***	0.0235**	–0.0490***	0.0411***
	(–11.94)	(3.76)	(–9.57)	(2.28)	(–6.46)	(4.09)
Large or secondary urban center	0.127***	0.446***	0.0896**	0.467***	0.107**	0.342***
	(3.72)	(7.13)	(–7.92)	(–8.49)	(2.20)	(4.33)
N	20065		20065		20065	

Note: t statistics in parentheses. Includes region dummies. *p<0.10; **p<0.05; ***p< .01.

TABLE 5A.5: Wage and Average Revenue Product of Labor Regression

	Corrected	OLS
Single worker OOME head	–59.57**	–3.900
	(–2.85)	(–1.06)
Multi-worker OOME head	68.47***	53.43***
	(10.86)	(11.58)
Received training for main job	76.56***	76.53***
	(8.41)	(14.62)
Age	1.318***	1.364***
	(12.02)	(15.26)
Male = 1	31.59***	29.87***
	(13.83)	(12.49)
Years of completed education	9.042***	8.819***
	(17.53)	(21.94)
Paid worker in OOME	–12.00***	–11.80***
	(–3.40)	(–3.51)
Father skilled salaried worker	–26.04***	–24.94***
	(–2.76)	(–4.49)
Father semi-skilled salaried worker	–28.95***	–27.30***
	(–2.64)	(–3.68)
Father salaried manual laborer	–38.35***	–35.65***
	(–3.89)	(–5.23)
Father head of own company	–29.70**	–30.05***
	(–2.17)	(–4.05)
Father worked on own (no salary)	–33.69***	–33.18***
	(–3.60)	(–6.06)
Father apprentice	–93.77***	–96.43**
	(–3.14)	(–2.26)
Father's years of education	1.175**	1.161**
	(1.93)	(2.42)
Mother's years of education	1.780***	1.727***
	(2.81)	(3.25)
Main job in public administration	113.6***	114.5***
	(14.16)	(21.96)
Agricultural laborer	–29.01***	–25.71***
	(–6.30)	(–4.44)
Large urban center	12.39***	12.85***
	(5.20)	(4.75)
Primary job in industry	–6.280	–3.569
	(–1.48)	(–0.65)
Primary job in services	–46.03***	–43.07***
	(–10.14)	(–7.68)
Primary job in trade	6.140	8.878
	(1.55)	(1.62)
_cons	14.58	11.17
	(1.30)	(1.00)
N	12455	12455

Note: t statistics in parentheses. Region dummies also included. *p<0.10; **p<0.05; ***p< .01.

TABLE 5A.6: Multivariate Tobit Estimation of Correlates of Paid Hours of Labor Employed (Last Month)

	Coefficient	t-stat
Owner received vocational training for main job	–88.7*	–1.93
Male	330.7***	7.68
Years of education	29.2***	5.13
Father head/boss	249.7	3.28
Father worked in agriculture	–114.7**	–2.54
Large urban center	146.4***	3.14
Industry	–259.3**	–2.55
Services	–423.8***	–3.82
Trade	–441.4***	–4.18
Constant	–1193.4***	–4.67
N	3722	

Note: Estimation includes region dummies and reasons for operating OOME.

TABLE 5A.7: Significant Determinants of OOME Registration

	Single-worker		Multiworker	
	Unregistered	Registered	Unregistered	Registered
Father semi-skilled salaried	0.322	–0.848**	0.346*	–0.734**
	(1.54)	(–2.12)	–1.84	(–2.10)
Father salaried manual laborer	0.539***	–0.396	0.494***	–0.0154
	(2.67)	(–1.07)	–2.72	(–0.05)
Father was head (no salary)	0.297	–0.0612	0.399**	0.238
	(1.46)	(–0.17)	–2.19	–0.76
Father worked on own (unsalaried)	0.337*	–0.187	0.358**	0.107
	(1.77)	(–0.56)	–2.11	–0.36
Father worked in agriculture	–0.350***	–0.351***	–0.358***	–0.367***
	(–5.51)	(–2.88)	(–6.09)	(–3.87)
Father worked in trading	0.0822	0.0774	0.146**	0.289***
	(1.13)	(0.58)	–2.16	–2.84
Father's years of education	–0.0120*	0.00250	–0.0133**	0.00258
	(–1.70)	(0.19)	(–2.05)	–0.25
Male = 1	–0.246***	0.226***	–0.1000***	0.256***
	(–7.44)	(3.35)	(–3.31)	–5.01
Age of head	0.0150***	0.0162***	0.0192***	0.0206***
	(13.29)	(7.12)	–18.44	–11.86
Years of education	–0.0543***	0.0235**	–0.0622***	0.0298***
	(–9.57)	(2.28)	(–11.94)	–3.76
Large or secondary urban center	0.0896**	0.467***	0.127***	0.446***
	(–7.92)	(–8.49)	(–8.87)	(–10.32)

Note: t statistics in parentheses. Region dummies also included. *p<0.10; **p<0.05; ***p< .01.

NOTES

1. These figures were calculated using survey information from Enquête Nationale sur les Objectifs Millenaire du Développement (ENSOMD) 2012.
2. The percentage of workers thus employed appears to have fallen, based on 1-2-3 surveys (nested household employment/labor force, and microenterprise, and poverty/consumption surveys), which are not comparable, between 1995 and 2010 (Nordman, Rakotomanana, and Rouboud 2012).
3. A key feature of these enterprises is that they are owner operated. Indeed, microenterprises operated by someone other than the owner are practically nonexistent in Madagascar.
4. They also find that firms that are privately owned, those that are incorporated, and those with more access to external financing are more innovative.
5. The United States presents a different case, however. There, once firm age is taken into account, firm size plays no significant role in job creation (Haltiwanger, Jarmin, and Miranda 2010).
6. For example, consumers in developing countries may prefer to purchase goods from a series of small merchants rather than from a supermarket with refrigeration, higher electricity, and advertising costs but offering greater convenience and quality or variety.
7. In some SSA countries, for example, microcredit is widely available to micro-entrepreneurs. In Togo, for example, there are over 140 microfinance institutions, which serve millions of clients, whereas in Madagascar, microcredit is more limited.
8. Unconditional probit estimation of heading an OOME in the respective sectors. In all cases except in industry, the coefficient on "male" was statistically significant at the 1 percent level.
9. This is nonetheless higher than the level of capital invested among urban SSA microenterprises that Grimm, Kruger, and Lay (2011) found (of 80 international dollars.)
10. We attempted to deduct depreciation of fixed assets to compute profits but were unable to obtain sensible results, and therefore utilize cash flow instead.
11. The returns to capital and labor are shown as separable here after testing whether cross-product terms were significant. They were significant only in the pooled sample, and for convenience we therefore assume separability.
12. Given that the survey lacked measures of physical units of output, as well as output prices and some input prices, including for unpaid labor, we could not estimate a production function per se. We therefore interpret our findings as the joint effects of average costs and different degrees of market power.
13. This procedure minimizes the squared sum of out-of-sample prediction error. Identification of the second stage equation is provided by the exclusion of several family history variables in that equation, as well by the nonlinearity of the relationships from the first stage.
14. In general, we lose up to 400 observations when using the correction step, due to the lack of family history variables. When OLS estimates are more precise and have a lower mean squared error, we report only these.
15. Tests of the significance of the difference in the coefficient on assets rejected the null of no difference at the 1 percent level.
16. The return to the head's time is higher than for other "unpaid" labor, and the return to unpaid labor is higher than that for paid labor, as one would expect, given the definition of returns as net cash flow within the period. Since there is typically some kind of noncash compensation for "unpaid" workers that is not captured in the firms' accounts, this does not mean that firms can expand their unpaid workforce without bearing any cost.
17. This is based upon a Mincer-style regression, with region, male, and urban indicator variables also variously included. Men are paid 37 percent more than females conditioning on these variables and years of schooling.
18. The potential for increasing profits by adjusting the head's and unpaid labor is typically not great for OOMEs. For single-worker firms, if one attributes a reservation wage to the head equal to the difference in the value marginal revenue product of paid and unpaid workers in multi-worker firms—wage of approximately 400 ariary per hour—the optimal level of the head's labor input in single-worker firms would be about 1.5 heads working full time. This is higher than the observed mean of 146 hours. However, since owners cannot bring in another owner without sharing profits, their labor allocation decision is not flexible enough to expand incrementally to 1.5 owners. Moreover, while using unpaid labor also raises profits, the firm's accounts do not fully capture the benefit flows to unpaid laborers out of profits, and typically the application of unpaid labor is constrained by the availability of family labor—not something the enterprise can optimize freely.
19. There is no correlation with the age of the firm, again suggesting that OOMEs generally decide on their scale of operation based on a set of constraints and opportunities at entry.
20. To identify technical increasing returns (decreasing average total costs), the elasticities of output with respect to inputs would be greater than one. Here, we have only measures of profits, rather than average total costs (ATC). Assuming that prices are constant, a positive elasticity implies that ATC are decreasing at that point of the factor use distribution.
21. If assets are lumpy, they could save up for them over a number of periods. Eventually, firm size would be limited either by market size or by exhausting increasing returns to scale.
22. If returns to enterprise inputs are decreasing, even if entrepreneurs are borrowing-constrained, without risk aversion they will invest and achieve profit maximization in a steady state. If entrepreneurs are risk averse, they will invest less, but will not be borrowing-constrained in a steady state (Osborne 2006). If there are increasing returns, however, this explanation is no longer adequate to explain underinvestment.
23. Even in this model, in some periods the entrepreneur will save positive assets, and at least one steady state will result with positive savings.
24. The precise ranges will depend upon parameters of the profit and utility functions, as well as the distribution of ε_{t+1}.
25. If there is a positive probability of a high draw of ε_{t+1} such as $x_{t+1} > \bar{x}$, then entrepreneurs may escape the poverty trap.
26. World Development Indicators reports interest rates of approximately 60 percent of that year, with 52 percent of this constituting the risk spread. However, we report World Bank staff estimates using central bank data.
27. In French, these are Registre du Commerce, Caisse Nationale de Protection Sociale, Patente, and Carte professionnelle.
28. For example, as is indicated by Doing Business *Employing Workers* data, Madagascar has the 10th highest minimum wage in the world as a percent of average labor productivity, at 0.9. Countries with a higher ratio are Burkina Faso, Haiti, Honduras, Kenya, Mozambique, Senegal, Sierra Leone, Togo, and Zimbabwe.

REFERENCES

Aghion, P., U. Akcigit, and P. Howitt. 2013. "What Do We Learn from Schumpeterian Growth Theory?" Schumpter Lecture presented at the Swedish Entrepreneurship Forum, Stockholm, January 2013. http://scholar.harvard.edu/files/aghion/files/what_do_we_learn_0.pdf?m=1361377935.

Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic. 2011. "Firm Innovation in Emerging Markets: The Role of Finance, Governance, and Competition." *Journal of Financial and Quantitative Analysis* 46 (6): 1545–1580.

Bento and Restuccia. 2014. "Misallocation, Establishment Size, and Productivity." University of Toronto Department of Economics Working Paper No. 517. July.

Das, M., W. Newey, and F. Vella. 2003. "Nonparametric Estimation of Sample Selection Models." *Review of Economic Studies*. 70: 33–58.

Greenwald, B., and J. Stiglitz. 1990. "Asymmetric Information and the New Theory of the Firm: Financial Constraints and Risk Behavior." *American Economic Review* 80 (2): 160–5.

Grimm, M., J. Kruger, and J. Lay. 2011. "Barriers to Entry and Returns to Capital in Informal Activities: Evidence from Sub-Saharan Africa." *Review of Income and Wealth* 57 (special issue, May).

Haltiwanger, J., R. Jarmin, and J. Miranda. 2010. "Who Creates Jobs? Small versus Large versus Young?" Working Paper 16300, National Bureau of Economic Research, Cambridge, MA.

INSTAT (National Institute of Statistics, Madagascar). 2013. *Enquête National sur l'Emploi et le Secteur Informel: ENEMPSI 2012*. Antananarivo: INSTAT.

Irinnews.org. 2013. "Microcredit helps small businesses buck the system in Madagascr." May 16. http://www.irinnews.org/news/2013/05/16/microcredit-helps-small-businesses-buck-system-madagascar.

LaPorta, R., and A. Shleifer. 2014. "Informality and Development." Working Paper No. 20205. National Bureau of Economic Research, Cambridge, MA.

Li, Y., and M. Rama. 2015. "Firm Dynamics, Productivity Growth, and Job Creation in Developing Countries: The Role of Micro- and Small Enterprises." *World Bank Research Observer* 30 (1): 3–38.

McKenzie, C. J., and C. Woodruff. 2006. "Do Entry Barriers Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises." *Economic Development and Cultural Change* 55(1): 3–42.

Montenegro, C., and H. Patrinos. 2014. "Comparable Estimates of Returns to Schooling around the World." Policy Research Working Paper 7020, World Bank, Washington, DC.

Neumark, D., B. Wall, and J. Zhang. 2009. "Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series." *Review of Economics and Statistics* 93 (1): 16–29.

Nordman, C., F. Rakotomanana, and F. Rouboud. 2012. "Informal versus Formal: A Panel Data Analysis of Earnings Gaps in Madagascar." DIAL Document de Travail, Antananarivo.

Osborne, T. 2006. "Credit and Risk in Rural Developing Economies." *Journal of Economic Dynamics and Control* 30: 541–68.

Poschke, M. 2014. "The Firm Size Distribution Across Countries and Skill Biased Technical Change in Entrepreneurial Technology." IZA Discussion Paper 7991, Institute for the Study of Labor, Bonn.

Rogerson, R., R. Shimer, and R. Wright. 2005. "Search-Theoretic Models of the Labor Market: A Survey." *Journal of Economic Literature* 43 (December): 959–88.

Sy and Diatta. 2014. "Jobs in Madagascar." Unpublished working paper. World Bank.

Van Biesebroeck, J. 2005. "Firm Size Matters: Growth and Productivity Growth in African Manufacturing." *Economic Development and Cultural Change* 53 (3): 545–583.

World Bank. 2015a. *Madagascar: Diagnostic Trade Integration Study—Levelling the Playing Field for Renewed and Inclusive Growth*. Internal draft for review. Washington, DC: World Bank.

World Bank. 2015b. *Madagascar: A Systematic Country Diagnostic*. Washington, DC: World Bank.