DOCUMENT DE TRAVAIL DT/2013-15

# Inputs, Gender Roles or Sharing Norms? Assessing the Gender Performance Gap Among Informal Entrepreneurs in Madagascar

*Christophe J. NORDMAN*
*Julia VAILLANT*

UMR DIAL 225
Place du Maréchal de Lattre de Tassigny 75775 • Paris Cedex 16 •Tél. (33) 01 44 05 45 42 • Fax (33) 01 44 05 45 45
• 4, rue d'Enghien • 75010 Paris • Tél. (33) 01 53 24 14 50 • Fax (33) 01 53 24 14 51
E-mail : dial@dial.prd.fr • Site : www.dial.prd.fr

# INPUTS, GENDER ROLES OR SHARING NORMS? ASSESSING THE GENDER PERFORMANCE GAP AMONG INFORMAL ENTREPRENEURS IN MADAGASCAR[1]

Christophe J. Nordman
IRD, UMR DIAL, 75010 Paris
PSL, Université Paris-Dauphine, LEDa, UMR DIAL, 75016 Paris, France
IZA
nordman@dial.prd.fr

Julia Vaillant
World Bank
IRD, UMR DIAL, 75010 Paris
vaillant@dial.prd.fr

**Document de travail UMR DIAL**
Octobre 2013

## Abstract

We use a representative sample of informal entrepreneurs in Madagascar to add new evidence on the magnitude of the gender performance gap. After controlling for business and entrepreneur characteristics, female-owned businesses exhibit a value added 28 percent lower than their male counterparts. Correcting for endogenous selection into informal self-employment raises the gap by 5 percentage points. We then investigate the role of sharing norms and gender-differentiated allocation of time within the household in the gender performance gap, by estimating their effect on the technical inefficiency of female and male entrepreneurs. Only male entrepreneurs seem subject to pressure to redistribute from the distant network. Our findings are consistent with situations where women working at home would essentially feel negatively the burden of their own community due to intense social norms and obligations in their workplace but also of domestic chores and responsibilities. We find evidence of females self-selecting themselves into industries in which they can combine market-oriented and domestic activities.

**Key words:** Gender, entrepreneurship, informal sector, sharing norms, household composition, Madagascar.

## Résumé

Nous utilisons un échantillon représentatif d'entrepreneurs informels à Antananarivo, Madagascar, pour mesurer et expliquer l'existence d'un écart de performance entre les unités de production informelles dirigées par des hommes et celles dirigées par des femmes. Une fois pris en compte les niveaux des facteurs de production, de capital humain, le secteur d'activité, l'année et la sélection endogène dans l'entreprenariat, l'écart de valeur ajoutée entre les entreprises féminines et masculines est d'environ 33%, au détriment des femmes. Nous étudions ensuite l'impact différencié des normes de partages au sein de la communauté et de la répartition des tâches au sein du ménage sur la capacité des hommes et des femmes entrepreneurs à atteindre leur frontière de production. Notre analyse suggère que seuls les entrepreneurs masculins sont sujets à la pression à la redistribution de la part du réseau distant. Pour les femmes, opérer une activité à domicile n'est pas un handicap en soi, mais cela agit plutôt comme un vecteur de transmission des effets négatifs des normes sociales et de répartition des tâches sur la gestion de l'entreprise. Nos résultats sont compatibles avec des situations dans lesquelles les femmes entrepreneures opérant une activité à domicile ressentiraient davantage le poids de leur propre communauté, sans doute à cause de normes de solidarité contraignantes, mais aussi à cause de leurs responsabilités domestiques.

**Mots Clés** : Genre, entreprenariat, secteur informel, normes de partage, allocation du temps au sein des ménages, Madagascar.

**JEL Code**: D13, D61, O12, J16.

[1] Acknowledgments :The authors thank Flore Gubert, Marc Raffinot, François-Charles Wolff, participants at the CSAE Conference 2012 in Oxford, at the *Journées Internationales du Développement* of Gretha/Gres 2012 in Bordeaux, at the 10th Midwest International Economic Development Conference 2013 in Madison, and at the XXIXth *Journées de l'Association Tiers-Monde* 2013 in Créteil for helpful comments and suggestions on earlier drafts. We are also indebted to William H. Greene for fruitful support in the use of his software NLOGIT 5.

## I. INTRODUCTION

In developing countries, women are disproportionately concentrated in low productivity activities, self-employment, and the informal sector. Many recent studies suggest that female-owned enterprises grow more slowly and exhibit lower profit and productivity levels than male-owned ones (Aterido and Hallward-Driemeier, 2011; Bardasi et al., 2011; De Mel et al., 2009; Mead and Liedholm, 1998; Nichter and Goldmark, 2009). A recent series of randomized controlled trials granting in-kind and cash transfers to male and female micro-entrepreneurs provide evidence on gender-differentiated returns to capital. In Sri Lanka the grant generates no increase in the profit of female-run businesses, unlike that of male-run ones (De Mel et al., 2008). Positive returns to capital for both men and women are found in Ghana, but for the latter the effect is significant only in the category of firms with initial profits above the median (Fafchamps et al., 2011). Performance differences can be explained by gender gaps in various dimensions, including women's generally lower level of human capital, lower stock of physical capital, higher concentration in low-performing activities, and also their lower ability to access and use financial services (Aterido et al., 2011; Mead and Liedholm, 1998; Bardasi et al., 2011).

So far, there has been limited work on business performance and returns to production factors differentiated by gender in the broad population of micro-entrepreneurs in Sub-Saharan Africa, and in particular in Madagascar, the setting of the present study. Even less is known about the reasons why women's incomes do not increase as much as those of men when they work as many hours and invest as much in their activity as their male counterparts.

Most empirical studies on the gender earnings gap in developing countries have looked at individual wages in the formal sector, ignoring thereby other possible sources of gender inequalities, in particular those occurring in informal self-employment, the largest category of workers in Sub-Saharan Africa. Using the latest waves of the Malagasy Enquêtes Prioritaires auprès des Ménages (EPM), Nordman et al. (2010) find that there is a rather small average gender earnings gap in Madagascar but that it is much higher in the non-farm self-employment sector. This differential is partly explained by a very unequal distribution of micro-firm attributes between female and male entrepreneurs, suggesting that access to physical capital may be an important source of earnings differential. For seven West African cities, Nordman et al. (2011) suggest that differences in sector allocation (public, formal private or informal sectors) contribute, on average, to one third of the gender earnings gaps measured on the broad population of workers. In fact, they also show that gender gaps are particularly wide in the informal sector, especially among the self-employed, for which a large part of the gap is explained by the characteristics of the micro-firm. In a recent study in Sub-Saharan Africa, Aterido et al. (2011) find that women are in fact not disadvantaged in terms of access to finance when key characteristics of the firm and the entrepreneur are controlled for. However, the available empirical evidence shows that part of the gender performance gap remains even after controlling for the industry, the human capital of the entrepreneur and the level of physical capital of the firm.

Against this background, the purpose of this contribution is twofold. Firstly, we add new evidence on the magnitude of the gender performance gap for informal entrepreneurs in the case of Madagascar, taking into account possible distributional and labor market selection effects. We use decompositions to assess the share of the gap that cannot be explained by differences in observable characteristics of the firm or the entrepreneur. Secondly, we explore the potential role that sharing norms and gender-specific allocation of time within the household play in this inequality. These factors may be responsible for technical inefficiencies in production, explaining why women under-perform.

The first hypothesis studied is the existence of adverse effects caused by redistributive pressure from the family and the kin group. In Madagascar, the prevailing solidarity system is called *fihavanana*. This word captures many dimensions: it suggests a moral obligation to consider others as relatives, translates into benevolent gestures, fraternity, mutual respect, cordiality, search of consensus, solidarity and sharing, in particular in difficult moments. An essential part of the *fihavanana* is the *famangiana*, or solidarity calls, to comfort or congratulate in case of happy (weddings, births) or unhappy events (deaths), and often include a monetary gift (Randrianja and Ellis, 2009; Razafindratsima, 2005; Wachsberger, 2009). The *fihavanana* is considered a specific Malagasy "way of life" and its importance is such that it is written in the preamble of the Constitution of Madagascar.[1] Many studies have found that family and kinship ties can be a vehicle for social contracts of mutual insurance in a context where markets for these goods and services do not exist. Fafchamps and Minten (2002) show for instance strong positive effects of social network capital on the performance of agricultural traders in Madagascar. However, family and kinship ties may become an important obstacle in the process of firm development if members of the kin system that achieve economic success are confronted with sharing obligations by less successful fellows (see e.g. Platteau, 2000; Hoff and Sen, 2006; Luke and Munshi, 2006; Grimm et al., 2013). This may imply remitting money, finding urban jobs or hosting them in the city home. Such prospects could adversely affect the incentives of an entrepreneur to pursue and develop his/her economic activity. A direct effect may also arise if part of the goods or services produced has to be shared, reducing thus the profit for a given level of inputs.

Given the central role of the *fihavanana*, studying the impact of sharing and solidarity norms in Madagascar seems particularly relevant. In this paper, we examine whether the pressure to redistribute is higher for women. Grimm et al. (2013), using similar data collected among informal firms in West Africa, find that looser ties with the family and kin group members who remained in the village of origin of migrant entrepreneurs are associated with higher capital and labor inputs, in particular for women.[2]

The second set of explanations examined is the impact of gender differentiated allocation of tasks within the household. Cultural norms defining the respective roles of women and men within the household and society may explain why female-run businesses tend to stay small and more subsistence-oriented. As traditionally the primary caretakers of children and responsible for domestic chores, women could choose self-employment because it offers flexible work arrangements. Gender-specific spending priorities also define the amount reinvested in the business, as females are known to devote a higher share of their earnings to the welfare of children (Grasmuck and Espinal,

2000; Duflo, 2003; Duflo and Udry, 2004). Fafchamps et al. (2011) find for Ghana that cash grants seem to be spent on household expenses and transfers rather than invested in the business. It has also been argued that women run their businesses in a subsistence-oriented manner to complement their husbands' income (Kevane and Wydick, 2001; Nichter and Goldmark, 2009). Women spend more time performing domestic tasks which can divert part of their most productive time away from market-oriented activities. We investigate whether gender differentiated allocation of time within the household can cause women to allocate their time sub-optimally to their firm.

Many women operate their business in their homes because it enables them to combine work and family activities. Besides, in some countries, home-based activities are often the only opportunity for women to generate income as the prevailing cultural norms prevent them from exercising an activity outside or far away from the home. Home-based activities carry several disadvantages, such as being far from input markets and clients (Ypeij, 2000), or rendering the firm resources visible and available to other family members, who may call on these resources if needed (Mead and Liedholm, 1998; Grasmuck and Espinal, 2000). As argued by De Mel et al. (2009), inefficiencies arising from a lack of cooperation within the household may be responsible for the lower returns to capital in women-run businesses found in Sri Lanka. They find evidence consistent with the spouse capturing part of the profit or working capital of the business. So far there has been limited empirical evidence on the impact of the location of the activity on the performance of these home-based "invisible entrepreneurs", partly due to the lack of proper data. In this study we use survey data representative of the informal sector in Antananarivo, which include a large share of home-based activities.

We use the 1-2-3 Surveys data collected among informal entrepreneurs in Antananarivo in 1995, 1998, 2001 and 2004, which are repeated cross-sections. This dataset and the variables used to proxy sharing norms and allocation of time within the household are presented in the next section. In Section 3, we measure the magnitude of the gender performance gap for informal entrepreneurs. To this end, we estimate production functions to assess gender differences in performance and factor returns. In a second set of regressions, we divide the sample into female- and male-owned informal firms, and run quantile regressions for each of these sub-samples to look at possible distributional effects. We also decompose the gender performance gap based on the previous regressions, to assess the extent to which the gap can be explained by differences in observable characteristics of female- and male-owned businesses and their respective owners. In Section 4, we study the impact of sharing norms and allocation of time within the household on the technical efficiency of informal businesses, defined as the distance between the actual output and the frontier of output, given the firm's and entrepreneur's characteristics. We make the hypothesis that these effects are gender-specific and to this end, we estimate stochastic frontier production functions and the determinants of inefficiency separately for female- and male-owned enterprises.

## II. DATA AND VARIABLES

### (a) Dataset and control variables

We use data from the *1-2-3 Surveys*, collected in Antananarivo between 1995 and 2004 (Rakotomanana et al., 2003; Razafindrakoto et al., 2009). These surveys were specifically designed to collect information among representative samples of informal firms. Phase 1 is a labor force survey, conducted every year since 1995, among 3,000 households. Principal and secondary activities of every member aged 10 years and over are recorded, including the type and status of the enterprise in which they work (formal/informal), making the establishment of a list of all informal firms headed by any member of a household possible (whether it is the main or the secondary activity). This list serves as the sampling frame for Phase 2, in which around 1,000 businesses randomly drawn from the Phase 1 listing are surveyed. The stratification scheme, by industry (nine) and type of firms (with or without wage workers), as well as an oversampling of the most atypical kind of firms (e.g. big manufacturing enterprises) make sure that the full heterogeneity of the informal sector is captured. Finally, Phase 3 is a household expenditure survey which interviews another representative sub-sample of households drawn from Phase 10. In this paper we mainly use data from Phase 2, which collects information on the characteristics of firms, such as the number and characteristics of all workers (including family helps), investment, expenditures for intermediate inputs, fees and taxes, sales and profits. Phase 2 has been conducted every three years since the start of the *1-2-3 Surveys*. Hence, we can use data on a representative sample of approximately 4,000 informal firms surveyed in 1995, 1998, 2001 or 2004.

We calculate the value added of each firm as the difference between sales and intermediary consumption. Intermediary consumption includes raw material and inventory purchases, rent and utilities, and other expenses. Value added thus includes capital income, all labor income (thus including the implicit cost of domestic unpaid work) and entrepreneurial profits. Capital is the total stock of capital measured at its actual replacement value.[3]

Let us now discuss how household characteristics are dealt with in the analysis. A large set of household variables is available in the data. To include the most meaningful variables and avoid collinearity problems in the econometric analysis, we first run a principal component analysis (PCA) over the following set of variables to define axes which summarize the household data and can be included in the regressions: number of female and male members by age category, size, dependency ratio, level of education (mean, variance, gap between head and spouse), mean potential experience, ethnic group, quality of housing and other wealth indicators, migratory status and the portfolio of activities of household members. We make use of the first five generated PCA axes as substitutes for household variables in the efficiency regressions in Section 4. By construction, these axes have indeed the advantage of being orthogonal to each other, therefore circumventing potential multicollinearity issues which might be important in the case of household characteristics. The pairwise correlation coefficients of the household characteristics are then used for the interpretation of the computed factors.[4]

### (b) Sharing norm proxies

In this paper we test the gender-differentiated effect of two different sets of variables on the efficiency of firms. The first set relates to the impact of sharing norms on the efficiency of the entrepreneur's activity. As argued in the introduction, the pressure from the family and the kin group to redistribute profits from one's activity can has

adverse incentive effects on the efficiency of the firm, that is, it might encourage the owner to allocate factors sub-optimally to production or lower his/her level of effort.

A first important proxy of social networks is the share of the population from the same ethnic group in the neighborhood in which a household resides. The census sampling zones ("segments") were used to define the neighborhood. This share is computed using the household questionnaire of the Phase 1 survey using population weights such that it exactly reflects the true share in the total population of the segment. As argued by Grimm et al. (2013), who also retain this approach, this measure of ethnic concentration is an indicator of the potential intensity of family and kinship ties, and more broadly community ties within the local network. The Merina are by far the most numerous ethnic group, representing almost 90 percent of the total population of Antananarivo and the second most important ethnic group is the Betsileo (4.5 percent).[5] Ethnic concentration is high but whether it is likely to act in favor of social pressure, or the reverse, is not clear. For instance, members of very small ethnic communities being surrounded by large ethnic groups are likely to attach more importance to social ties and hence maintain and/or reinforce them.

Secondly, we use a measure of the distance to the district of origin as a proxy for the potential intensity of kinship ties within the distant network. We assume that the further away a person lives from his/her district of origin, the more difficult it is for the family to observe the entrepreneur's activity and thus exert redistributive pressure. Phase 1 includes a question on whether the individual has always been living in Antananarivo and, if not, what district in the country does he/she come from. For this variable, we relied on two proxies: the geographical distance calculated in kilometers (computed from Antananarivo to the district of origin using geographical maps) and the travel time to this district. The latter proved to be more efficient in approaching a (time and monetary) cost to keep in touch with the remote family, in a context where roads could be in very different states.[6] The travel time variable is included in quadratic form to take into account non-linearities in its effect. For non-migrant entrepreneurs, it is equal to zero. The sample used in the efficiency analysis counts 791 migrants, roughly 27 percent of all observations. Estimating the efficiency model on the sub-sample of migrants yielded robust results.

Thirdly, we build a variable indicating whether the entrepreneur provided free services to the community in the previous week. We hypothesize that this is a measure of links between an entrepreneur and his/her community, and he/she might be engaged in some sort of reciprocity agreement with members of his/her community. This is done thanks to a question on time use in the previous week included in the labor force survey (Phase 1), which asks for the number of hours spent performing various tasks, such as providing free services to the community, performing domestic tasks, building the house or studying.

These three variables are believed to be proxies of the redistributive pressure that an entrepreneur may potentially feel. A more direct measure of such pressure could be monetary or in-kind transfers given and received by the entrepreneur. However this poses a data problem, as transfers are available in Phase 3 of the *1-2-3 Surveys*, which is only carried out on a small sub-sample of informal businesses surveyed in Phase 2. In addition, transfers are clearly endogenous as they are a share of a household's income which itself depends on the IPU's value added. Therefore

including transfers directly into production functions could potentially bias the estimates. It can be of interest, however, to check whether our proxies of redistributive pressure are indeed correlated with monetary transfers given and received by an entrepreneur. Using Phase 3 matched with Phase 2 data, we regressed transfers (in log and in Malagasy Francs) given and received on our redistributive pressure proxies. Controlling for total expenditures and household characteristics, we found that each of our proxies was significantly correlated with transfers in at least one of the models.[7] Similar proxies proved to be correlated with transfers in several other studies on social networks (Comola and Fafchamps, 2010; Nguyen and Nordman, 2012; Grimm et al., 2013).

(c) Time allocation within the household

The second set of variables pertains to the role the household plays in the productive process. Gender norms in the allocation of tasks within the household may have an important impact on the efficiency of the firm.

The question on time use detailed above is used to obtain a variable measuring the number of hours spent by the entrepreneur in the previous week doing domestic tasks. The time devoted to going to the market and fetching wood and water, a separate category in the questionnaire, is added to domestic tasks to obtain a more complete measure of domestic time use.[8] We hypothesize that spending a high number of hours on household chores is energy consuming and interferes with the productive activity of the entrepreneur. This should be particularly true for women who devote the most time to these domestic activities.

The composition of the household in terms of active and inactive members may also have a gender- differentiated impact on entrepreneurs. The amount of hours they spend on their activity can be limited by the presence in the household of little children in need of constant care and attention, usually provided by the mother. The performance of the activity itself may be less efficient if the entrepreneur is interrupted often by the need to look after children. Having elderly and perhaps dependent individuals in the household can play a similar adverse role. We create two variables using the household questionnaire of the Phase 1 survey: the number of children aged four years old and younger in the household, and the number of household members aged 65 and older, excluding the entrepreneur and his/her spouse.

The effect of the composition of the household might also relate to the redistributive pressure hypothesis. For example, it is possible that the entrepreneur feels a pressure to share with other members of the household who are not working, in this case, the elderly. Another type of solidarity mechanism, child fostering, is often used by poorer, often rural households, who send their children to live with urban relatives (Mahieu, 1990). A negative effect of the presence of several young children in a household on the efficiency of production could then reflect another effect of sharing norms. Lastly, the labor force module in Phase 2 (firm-level questionnaire) provides information on the composition of the workforce in the firm, including the relationship of each worker to the owner of the firm. We include a variable measuring the share of family labor in total labor, equal to the ratio of the number of hours worked by family workers over the total number of hours worked in the firm.

(d) Description of female- and male-owned firms

The data show that the value added of female-owned firms is lower than that of male-owned firms (Table 1). This is true for every year and across the entire distribution of value added. However there are changes over time. An analysis of the dynamics of the gap and the respective roles of differences in characteristics and differences in returns along the distribution is provided in Section 3.3

Table 1 shows the mean average characteristics of IPUs, their owners, and their respective households, by sex of the entrepreneur. In terms of factor endowments, female-owned firms operate with much less physical capital than male-owned firms: their stock of capital being on average less than half that of their male counterparts. The level of labor, measured by the number of hours worked in the previous month is also inferior, especially the number of hours worked by family members and hired workers. The labor input of the owner is lower for females but of comparable magnitude.

A number of firm characteristics seem to be gender specific. The distribution of firms across sectors of activity is more concentrated for female entrepreneurs than for males. Two sectors only (out of nine) concentrate half of the female entrepreneurs in the sample: textile and trade of primary goods. Women are virtually absent from the construction and transport sectors, which are predominantly male activities. Female-owned firms are smaller on average, and more frequently operated solely by the owner. With a lower percentage of labor provided by family workers, and their partner working less frequently in the firm (if married), female entrepreneurs seem to benefit less from the help of their family to operate their activity than their male counterparts. On the other hand, their activity is more embedded in the household as it is located in the home of the entrepreneur for a third of female firms, against 17 percent of male firms.

The schooling achievement of female and male entrepreneurs is similar, although female entrepreneurs have less frequently completed high school or started higher education. Female owners are also slightly younger and their experience, measured by the number of years the entrepreneur has practiced her main activity, is 1.5 years lower for females than for males. The marital status of the entrepreneur is also gender specific, as more male entrepreneurs are married than females. In the same vein, female entrepreneurs belong to households with a lower dependency ratio[9] than males, and with a smaller number of children under five years old. As expected, female entrepreneurs devote far more hours to domestic tasks, fetching wood and water or going to the market than males. For example, women spent 13.6 hours performing domestic tasks in the week prior to the interview, whereas men had spent only 3.7 hours.

Female entrepreneurs belong more frequently to households in which there is a member (who is not the owner of the considered IPU) working in the public sector or in the private formal sector. This suggests that informal firms owned by women act as a way of diversifying sources of income in households in which there is already a steady source of income.

## III. MEASUREMENT OF THE GENDER PERFORMANCE GAP

### (a) Effect of gender on value added and factor returns

#### (i) Econometric model

We first estimate production functions to assess the existence and magnitude of gender differences in firm performance and factor returns. Let $Y_j$ be the monthly value added of IPU j. We estimate the following Cobb-Douglas production function:

$$lnY_j = Inputs_j\beta_1 + Owner_j\beta_2 + IPU_j\beta_3 + Year_j\beta_4 + \varepsilon_j \qquad (1)$$

where $Inputs_j$ is a vector of input variables ($1, lnK_j, lnLo_j, lnLf_j, lnLh_j, D_j$). $K_j$ stands for physical capital of informal business $j$. To take into account the heterogeneity in the type of workers, labor is split into the number of hours worked monthly by the owner ($Lo_j$), family workers ($Lf_j$), and hired workers ($Lh_j$). Family workers have a family relationship with the IPU owner and are precisely identified thanks to the firm level questionnaire (Phase 2). A number of IPUs do not use all types of labor or capital, but excluding them from the regression or replacing the log value with an arbitrarily small positive value can bias the estimates. To avoid such a bias, for a given input that is not used by the IPU, we follow Battese (1997) and set the log-value of the input to zero, while controlling with dummy variables equal to one if the level of the input is positive ($D_j$ vector). $Owner_j$ is a vector of characteristics of the owner of IPU $j$, including his/her sex, education, experience, age, age squared and marital status. $IPU_j$ is a vector of characteristics of the IPU, namely, the time in business and eight sector of activity dummies. $Year_j$ is a vector of year dummies. We first estimate equation (1) without and with the entrepreneur's human capital variables.

In order to check the sensitivity of the estimated determinants of firm performance to the presence of unobservables at the household level, we then add household fixed effects. The drawback of this method is that we have to reduce the sample to the group of firms belonging to households for which data are available on more than one firm (25 percent of the initial sample). This is why we use this technique in only one set of regressions so as to avoid reducing drastically the sample size in the rest of the analysis. By pooling the data across sexes, we constrain the effects of the determinants of firm performance to be identical for males and females. We relax this assumption and check whether these factor returns differ across gender by estimating production functions with a set of interactions with the gender dummy.

Let us now discuss a few econometric problems common to the estimation of production functions. Although the available data limit what can be done to correct some of these issues, we attempt to present how they might bias the coefficients in the production function and how we circumvent these issues when possible. A first problem is the simultaneity of the level of observed inputs and output. Labor and capital are chosen by the entrepreneur and may be correlated with an unobserved productivity shock or an unobserved input, such as the managerial ability of the owner (Ackerberg et al., 2007; Aguirregabiria, 2009). The estimated coefficients of the input variables can therefore be biased upwards, in particular labor which is more flexible than capital and thus more easily adjusted following a shock. In the same vein, an upward bias will be caused by reverse causality between

the level of profit and capital. There a several ways to correct for this bias, such as instrumenting with input prices or using panel data.[10] With the cross-sectional data at hand, we can only use the fact that we have several firms belonging to the same household to control for part of the ability bias. This would mitigate part of the bias under the assumption that entrepreneurs belonging to the same household are more similar to each other than those belonging to separate households.

A second common problem in the estimation of production functions is endogenous exit which introduces a selection bias. In a cross-section, as we move up the distribution of firm age, the only firms we observe are survivors. As smaller firms are known to be more vulnerable than larger ones, surviving small firms are likely to be selected and have high levels of outputs. This would tend to bias downwards the capital stock coefficient (Ackerberg et al., 2007; Aguirregabiria, 2009). As female-owned firms are smaller, grow slower, and are less likely to survive the year than male-owned firms, the positive selection is likely to be more pronounced for surviving female-owned businesses. If this is the case, the downward bias on the capital coefficient will be stronger for females than for males and the gap in estimated returns is likely to be an upper bound estimate of the real gap.

Finally, estimating production functions using selected populations of entrepreneurs raises concerns over possible sample selection biases. Strictly speaking, there are two sources of selectivity involved. One arises from the fact that the value added of the self-employed is only observed when they earned an income and their activity was profitable, which is not the case of everyone. The second comes from the selective decision to engage in self-employment in the informal sector rather than in other employment states. One of the ways to account for sample selectivity related to labor market participation and sector choice is to use a generalization of Heckman's procedure (Lee, 1983) that takes into account the possible effect on earnings of endogenous selection in different employment states. In the first stage, multinomial logit models are used to compute the correction terms from the predicted probabilities of individual *i* being in employment state *j*: out of employment (inactivity and unemployment), formal wage employment, informal wage employment, informal self-employment. The generalized forms of the inverse Mill's ratios are then introduced into the value added equation for each employment state *j* and yield consistent estimators of the determinants of value added. Lee's method has been criticized, for it relies upon a strong assumption regarding the joint distribution of error terms of the equations of interest (Bourguignon, Fournier and Gurgand, 2007). However, the we tried existing alternative methods, such as Dubin and McFadden's or Dahl's, which did not provide very different results. Therefore we chose Lee's correction method which has the advantage of providing an easier interpretation of the correction terms.[11]

In Heckman's and Lee's procedures, identification is achieved using exclusion restrictions, i.e. by the inclusion of additional regressors in the first stage selection equations. To correct for endogenous selection of individuals into informal self-employment, we use Phase 1 of the *1-2-3 Surveys* (labor force survey), which provides a representative sample of individuals that are inactive or unemployed, formal sector wage workers, informal sector wage workers or informal self-employed. The first stage of the

selection model is thus a multinomial logit modeling the probability of being in these various employment states (the reference category being the out of employment individuals). The exclusion restriction we use is the status in employment of the father of the individual, whether he was a wage worker, an informal self-employed (no wage workers) or an informal enterprise owner (with wage workers. The idea is that these variables could impact the allocation of workers into the various employment states, due for instance to one's inclination to take up the same type of activity as that of their father, thanks to social norms, inheritance rules, and/or family tradition. There is however no clear reason to believe that, in Madagascar, the father's occupation at an early age of his child would provide the latter's business with any direct economic advantage later on (although it might do at the business' startup). Pasquier-Doumer (2013) finds in West-African capitals that informal entrepreneurs are not advantaged in terms of performance when their father is also informally self-employed, except when his activity is in the same sector. To make sure therefore that our identifying variables are exogenous, we interact the father's employment state with a variable indicating if the father's activity was agriculture or not. As all the informal entrepreneurs in our sample are non-agricultural (by sampling) this ensures that having a self-employed father in agriculture is exogenous to the performance of the studied IPU. We tested the appropriateness of this identification strategy using Wald tests of joint significance of the identifying variables in the first stage. These tests of joint significance of the instruments never rejected the null at the 1 percent level. This was even more so when the father's occupations were interacted with a dummy indicating his previous agricultural attachment, hence a possible previous migration episode of the family and/or the young entrepreneur from a rural area to the capital city.

Bearing in mind the methodological controversies surrounding the choice of identifying variables in general, we also report results from uncorrected production functions when possible. This will also help provide comparability with existing studies.

(ii) Results

Results are shown in Table 2. The first two columns of Table 2 show that labor and physical capital are positively and significantly associated with value added. Splitting the labor factor into three components (owner, family and hired workers) reveals that the owner's labor has slightly higher returns than hired labor, and that family labor is the least productive. To our knowledge, there are very little studies that have specifically addressed the nature and effect of family versus hired labor in firm performance. The small existing literature relates essentially to farm businesses. *A priori*, one might expect family and hired labor to play differently on business performance, because they may have different compositions of male and female, adult versus child, and skilled versus unskilled labor.[12] In addition, the composition of tasks performed by both types of labor should also be considered, and may

contradict the common assumption that hired labor is necessarily more productive than family labor.[13] Since family workers may perform management and supervisory duties (particularly the household head), their work may have larger effects on output than that of hired workers, who may only perform manual tasks. The performance of managerial and supervisory tasks by family members may then reduce the substitutability between family and hired labor. All this would explain why it is not clear that with family labor the entrepreneur would face supervision advantages, which were thought to come as family labor would share the benefits of work. Here, for Madagascar, we find that the hours of family labor are less productive than both the hired labor's and the entrepreneur's hours. This result is even reinforced once we control for unobservables at the household level using fixed effects (Column 4).[14]

Turning to our variable of interest, the sex dummy, we verify that firms owned by women perform less well than those owned by males, even after controlling for the level of inputs and the human capital of the entrepreneur. Column (2) shows a level of value added that is about 28 percent lower in female-owned firms than in male-owned firms. The decompositions of the raw gender value added gap in the next section will precisely tell us how much is actually explained by these observable characteristics and what proportion of this gap remains unexplained in our model.[15]

The regression in Column (4) includes a household fixed effect and is run on the sample of 984 IPUs for which information on another business belonging to the same household is available. There are thus 464 households in this sample. Compared to the OLSregression shown in Column (3) which is run on the same sample, these results reveal much lower returns to capital within the household. Within-household capital returns hint at the level of efficiency of the allocation of capital among the various activities of the household. A Pareto-optimal allocation would imply zero returns to capital, because capital should be allocated where returns are positive, and returns being decreasing, this should equalize returns inside the household and bring them to zero.[16] In our case, lower returns to capital found in the household fixed effect regression suggest a rather efficient intra-household allocation of capital. The effect of the entrepreneur's education on the firm performance becomes non-significant, maybe due to the smaller variance of schooling achievement within a household (several firms belonging to a given household may have the same owner).

As shown in Column (3), the female dummy has a much lower effect when a simple OLS regression is run on this sub-sample, which may be due to particular characteristics of this group of businesses. When the household fixed effect is added (Column 4), the gender gap increases and IPUs owned by females are still roughly 22 percent less productive than those owned by males of the same household. With inputs, sector and year held equal, the performance gap between a female and a male belonging to the same household is larger than the gap between a female and a male in two different households. Enterprises run by men may be favored over those run by female members of the household, but not necessarily in terms of access to resources, such as physical capital, as this result is holding all inputs equal. This result rather suggests that the effort put into the different activities is different, for example that productive time within the household is allocated to the advantage of the male-run

enterprise rather than the female-run one. As mentioned in the introduction, wives often run a business to complement their husbands' income, and in a manner which enables them to combine domestic activities and market-oriented activities. Such effects may not be directly observed using production functions, because they impact the technical efficiency of the productive process, that is, whether inputs are combined in the most optimal way or not. We tackle this question in Section 4.

The regression in Column (5) includes interactions of the explanatory variables with the female ownership dummy. We find that returns to capital in female-owned firms are significantly lower than that of male-owned firms. However, the hours of labor of female entrepreneurs have a return that is non-significantly lower than that of males. The same result holds for family and hired labor, exhibiting returns that are non-significantly lower in female-owned firms. We go back to the issue of relative productivity of the various factors in the next section. Finally, the effect on value added of having completed middle school (compared to having no schooling) is higher for women than for men, while returns to primary and higher education are similar.[17] We perform a F-test of joint significance of the interacted terms which rejects the hypothesis that the crossed gender effects are jointly zero at the one percent level (see bottom of Table 2). This means that the analysis of the determinants of firm performance should be carried out separately for each sex.

Finally, we run a regression including the Lee's selectivity correction term (Column 6). The estimated coefficient of the selection term is significant at the 5 percent level with a negative sign. This means that the mechanism of allocation across the four employment states is not random and affects value added significantly. Informal self-employment participation is associated with unobserved characteristics that are negatively correlated to business performance. Would sample selectivity not be accounted for, OLS estimates could then yield biased estimates of the returns to the observed determinants of value added. Interestingly, however, the results show that this correction only affects the coefficients of the owner's human capital, not those of the firm inputs. The returns to production factors, capital and labor in particular, are indeed essentially unchanged from Column (2) to Column (6). By contrast, the returns to the owner's age and education are clearly refined, showing that the returns were overestimated without introduction of the selectivity correction. For this reason, we shall pursue our analysis considering the selectivity-corrected production function in addition to OLS estimates.

### (b) Distributional effects of gender on value added and factor returns

In a second set of regressions, we divide the sample into female- and male-owned informal businesses (Table 3). We also want to shed light on potential distributional effects of the determinants of firm performance and so rely on quantile regressions for each of these sub-samples. The goal is to evaluate whether differences in performance and returns between men and women are different along the conditional distribution of the value added (Tables 4 and 5). Here the selection bias is corrected for females and males separately, that is, the selection process (multinomial logit) is not assumed to be the same for women and men anymore. To correct for sample selection in the quantile regression, Buchinsky (1998) suggests using a non-parametric

estimator of the selection term in the first stage. However, with our large sample sizes and number of covariates, we chose to follow Bollinger et al. (2011) and to include the (parametric) selection term computed from Lee's method described above as an additional regressor in the second stage).

Before we look at returns to factors, let us first briefly comment on the coefficients of the characteristics of the owner and the firm. The effect of having completed middle school compared to having no schooling is much higher for women than for men (Columns 1 and 3, Table 3), in particular when we correct for selectivity into informal self-employment (Columns 2 and 4). The effects of low levels of schooling for men even become non-significant with the introduction of the selectivity correction, while they remain relatively stable for women. This effects of low and middle levels of schooling for female entrepreneurs exist notably in the upper part of the conditional distribution of value added (Table 5)[18], from the first conditional quartile for primary education, and from the conditional median for middle school. While higher education provides benefits to male entrepreneurs almost all along the conditional distribution of business performance (Table 4), the effect of higher education is larger in magnitude for their female counterparts, especially at the lowest and highest ends of the distribution of value added. Interestingly, the professional experience of the owner is neither significant for males nor for females all along the distribution. The marital status exhibits a gender specific effect above the fourth quartile, as being married has a positive impact on the valued added of males, but no significant effect for females. This seems to suggest that in well-performing firms run by men, they benefit from the help of their wife, either in running the firm or in performing domestic tasks that directly help the entrepreneur in his managing the business. Conversely, the presence of a husband does not help female entrepreneurs.[19]

Turning now to returns to factors, we obtain increasing returns to capital for both males and females along the conditional distribution until about the median: from 11 percent around the first decile to about 16 percent at the median for men (respectively, roughly 8 percent to 11 percent for women). Then, the elasticities remain remarkably stable. All along the distribution, we confirm that the returns to capital are always lower for females than for males. By contrast, returns to labor of the owner are decreasing along the distribution for both women and men and these returns are always higher for females. For male entrepreneurs, family labor shows increasing returns and hired labor shows decreasing returns. In female-owned firms the picture is less clear. Hired labor has a much lower productivity than in male- owned firms, except in the highest decile, where it is significant and much higher than for males.

Looking only at returns to factors at various conditional quantiles may be misleading, as the intensity of each type of factor may be very different in male- and female-owned firms across the distribution. To get a better grasp of gender-specific effects in the substitutability between factors, we calculate the Technical Rate of Substitution (TRS), defined as the rate at which the use of one input $x_1$ has to increase to maintain the same level of output y when the other input $x_2$ is decreased by one unit. The general formula for the TRS is:

$$TRS_{1,2} = \frac{dy/dx_1}{dy/dx_2} \quad (2)$$

We calculate the TRS between capital and owner's labor at the mean, using results of the estimated production function defined by equation (1) and shown in Tables 3 to 5:

$$TRS_{capital,owner\ labor} = \frac{\widehat{\beta_{1k}}}{\widehat{\beta_{1Lo}}} * \frac{\overline{L_o}}{\overline{K}} \quad (3)$$

$\widehat{\beta_{1k}}$ and $\widehat{\beta_{1Lo}}$ are the estimated elasticities of value added with respect to capital and owner's labor, $L_O$ is the mean number of hours of labor provided by the owner and $\overline{K}$ is the average stock of capital in the sample. The TRS is calculated for male and female entrepreneurs separately, and is also given at the 10th, 25th, 50th, 75th and 90th percentiles, using the estimated elasticities found in the selectivity-corrected quantile regressions for males and females respectively (Tables 4 and 5). The TRS between the owner's labor and family labor, owner's labor and hired labor, and family labor and hired labor are calculated analogously. Results are shown in Table 6.

Table 6 shows that while, at the mean, the TRS in male- and female-owned firms are roughly the same, there are strong distributional effects that are worth discussing. Let us first comment on the substitutability between the entrepreneur's labor and family labor. At the bottom of the distribution (first decile), a decrease in one hour of the entrepreneur's labor requires an increase in family labor of 24 minutes (0.22 hours) in female-owned firms, and roughly one hour in male-owned firms, to keep the level of output constant. Above the fourth quartile the effect is completely reversed, well-performing female entrepreneurs need three times as much family labor to compensate a decrease in their labor than male entrepreneurs do, and twice as much in the top decile of the distribution. This could be explained by gender specificities in the type of activity and entrepreneurship found at the bottom and the top of the distribution of value added. Low-performing female entrepreneurs (bottom quantiles) accomplish tasks requiring little skills, and they can easily be replaced by family workers. At the top of the distribution, they perform more complex tasks which require a more intensive use of family workers to compensate a decrease in their work. Conversely, in low-performing firms owned by males, their work is already rather specialized and less easily replaced by family laborers. One can think of "masculine" tasks such as construction, or pushing a heavy cart, less easily performed by women or children for example. Table 6 hints at a potential constraint to the growth and success of female-owned firms. As her firm grows bigger, the entrepreneur's own work becomes more and more necessary because she is less replaceable. However we know that women are time constrained by domestic tasks they have to perform (see for example Table 1).

The TRS of the owner's labor for hired labor is increasing along the distribution for both sexes, meaning that entrepreneurs may be constrained in their growth by insufficient labor input if labor markets are imperfect. However, male entrepreneurs may be able to compensate using family labor (if available), which is not the case of females. Family and hired labor are however less and less substitutable as we move up the distribution of conditional value added: family labor is less replaceable in well-performing firms. This may be because family workers often take

up supervisory duties in large firms. Finally, capital is less easily replaced by the labor of the owner in male-owned firms, which could be due to a generally higher capitalistic intensity of their firms.[20]

### (c) Decomposition of the gender performance gap

After having highlighted genuine differences in the determinants of firm performance across gender, it is now necessary to estimate the extent to which the gender performance gap is due to differences in factor endowments, firm characteristics, entrepreneurs' human capital and demographics, versus differences between sexes in the returns to these attributes. This is the purpose of the Neumark and Machado-Mata decompositions presented in this section (Neumark, 1988; Machado and Mata, 2005). The decompositions, based on the previous regressions, tell us how much of the performance gap between female- and male-owned informal firms can be explained by differences in observable characteristics. The unexplained share of the decomposition corresponds to differences in returns to factors. In our framework, the unexplained share of the gap could be due to discrimination against female entrepreneurs stemming from consumers who prefer, for example, to buy products from males than females.

The gender gap literature, mainly for developed countries, has stressed the possible existence of a varying wage gap along the distribution of income (Albrecht et al., 2003; De la Rica et al., 2008; Jellal et al., 2008; Nordman and Wolff, 2009a). This literature highlights greater gaps at the upper end of the earnings distribution, the so-called glass ceiling hypothesis according to which women face more difficulties than men in reaching top positions within the firm and thus in benefiting from high wages. Evidence on this for developing countries is rather scarce but, in the case of Madagascar, the existing results show a gender wage pattern rather inconsistent with the glass ceiling phenomenon (Nordman and Wolff, 2009b). However, the studies tackling this issue looked mainly at individual wages in the formal sector, ignoring thereby other possible sources of gender inequalities along the distribution, in particular those occurring for informal self-employed workers. We seek evidence of such effects by performing the decomposition at each quantile of the performance distribution following the method developed by Machado and Mata (2005) for the analysis of changes in wage distributions.[21] We control for selectivity effects using as the dependent variable the value added net of the impact of the selection, that is $\overline{Y}_J - \widehat{\theta}_J\widehat{\lambda}_J$ with $\widehat{\theta}_J$ the estimated coefficient of the $\widehat{\lambda}_J$ selectivity correction term (Neuman and Oaxaca, 2004).

Looking at the pooled data, 42.4 percent of the raw gap in value added is explained by differences in the level of capital and labor used to operate the firm (Table 7, Columns 1 and 2). Adding the sector of activity and age of firm brings this share up to almost 70 percent. As shown by the descriptive statistics, the distribution of activities across firms is very gender specific, as females are more concentrated in a few sectors, while male entrepreneurs are more evenly distributed. The human capital of the owner increases the explained share of the gap by almost 6 percentage points. As a consequence, a quarter of the gap remains unexplained by differences in mean observable characteristics.

Looking at Columns (3) and (4), where selection into informal self-employment is accounted for in the calculation and decomposition of the gap (which reduces the raw difference to 61.4 percent), we observe that a greater share of the gap can actually be explained by production factors only, with an explained share of the gap reaching now almost 51 percent. Hence, more than 8 percent of the gender gap in business performances may be attributed to selectivity into informal self-employment across the sexes. The subsequent introduction of the IPUs' and owners' characteristics brings the explained shares of the gap to nearly 78 and 87 percent, respectively. In a nutshell, we are still left with more than 15 percent of the gap that remains unexplained by differences in mean observable characteristics, and gender-specific employment selection.

The quantile decompositions exhibit interesting distributional and temporal effects. Looking at the quantile decompositions pooled across years, we see that the share of the gap that is unexplained increases as we move up the distribution of value added (Table 7, Columns 2 and 4). This result still holds once selectivity effects are accounted for. Interestingly, at the bottom of the distribution (around the first conditional decile), the entire gap is explained by the IPUs' and owners' observed characteristics, together with gender-specific selection effects (Column 4). By contrast, the unexplained share of the gap widens continuously until around the fourth conditional quartile, reaching there almost 39 percent. Then, it diminishes again slightly to reach about one third of the raw gap. These findings are clearly in accordance with the hypothesis of the presence of a glass ceiling phenomenon among self-employed informal workers, which is not observed among formal wage workers, as shown in Nordman and Wolff (2009b). There are probably additional characteristics (either at the household or community levels, such as social networks), which remain uncontrolled for in our models, that may be more relevant to explain why men perform better than women at the upper end of the distribution of value added.

The general trend over the decade is a widening of the gender performance gap, except in 2001 where it sharply drops (see also Figure 1). In 2004, the gap is less well explained by observable characteristics than in 1995, an explained share that actually continuously diminishes across years from 85 percent in 1995 to 77 percent in 2004. The quantile decomposition shows that the gap in fact increased only at the bottom of the distribution, where the explained share of the gap also diminished.

The most salient feature of these decompositions is the sharp drop in the raw gap all along the distribution (except at the top decile) between 1998 and 2001. On plausible explanation would be that the worse performing females left the informal sector during that period of sustained macroeconomic growth, fostered among other things, by the development of Export Processing Zones (EPZ) in Madagascar. As discussed in Vaillant et al. (2011), a large number of women were hired in EPZ industries, which expanded very quickly during that period. This then potentially modified the informal firm allocation across gender and sectors.

This explanation is consistent with the evolution in 2001-2004, showing a strong increase in the gap at the bottom of the distribution. The share of the gap explained by factor inputs decreased during that period, while the share of the gap explained by other characteristics of the firm and the entrepreneur increased. This may indicate that, after the 2002 crisis, many poorly skilled women undertook informal activities in order to provide additional

financial resources to their household. The data show that the proportion of female entrepreneurs increased between 2001 and 2004, after having steadily decreased in the previous years. In particular, many unskilled females lost their jobs in Export Processing Zones industries due to the 2002 crisis.

In addition, between 2001 and 2004, at the high end of the distribution, the share of the explained gap exhibits a dramatic increase. One possible explanation is that, between these two years, there was a change in the skill composition of successful female entrepreneurs: perhaps after the 2002 crisis, many skilled women who previously ran their own successful business had to give up their activity to take care of other declining household businesses managed by their spouse. This may explain why successful women in 2004 are not the same as those in 2001, the latter benefiting from greater human capital endowment than the former.

Although the temporal analysis is interesting in itself, in the remainder of the paper we will not take it any further, because we wish to study other dimensions, and sample size would not allow us to disaggregate by year in addition to the other dimensions we choose to study here.

## IV. GENDER AND FIRM EFFICIENCY

### (a) Concept of technical efficiency and empirical strategy

In the previous sections, we found a 75 percent gap between female- and male-owned informal businesses, and that a quarter (or 15 percent, depending on whether we account for selection into informal self-employment or not) of this gap remained unexplained by observable characteristics such as the level of inputs, human capital of the entrepreneur or sectoral distribution. In this section we investigate whether sharing norms and within-household allocation of time may have gender-differentiated effects that could explain why female under-perform compared to males. We explore here how these effects impact the technical efficiency of informal businesses.[22] A producer is technically efficient if increasing his output requires an increase in at least one input or if he could not produce the same output when reducing one input.[23] Technical inefficiency is thus measured as the distance between a firm's actual output and its potential output given its level of inputs and characteristics.[24]

Sharing norms can create a sub-optimal use of inputs, if the entrepreneur feels that he/she will have to share part of his/her profits, he/she might work less intensively, or give less efficient tasks to his/her workers. A constrained use of family labor, in excess of the optimal needs (given capital stock), could create a level of productivity inferior to its optimal level. Involuntary inefficiency effects can also arise when part of the goods or services produced by the business is consumed by members of the social network, thereby reducing the (sold) value added, for a given level of inputs. The allocation of tasks within a household and its composition can also affect efficiency by distracting the producer and diverting the most productive time away from market-oriented production.

We follow the model developed by Battese and Coelli (1995) to estimate simultaneously the stochastic frontier production function and the determinants of efficiency using maximum likelihood. The stochastic frontier production model is defined as:

$$lnY_j = Inputs_j\beta_1 + Sector_j\beta_2 + Year_j\beta_3 + (v_j - u_j), \quad (4)$$

where $Y_j$ is the valued added of firm j. *Inputs* and *Year* are defined as in equation (1). Sector is a vector of indicators of the sector of IPU *j*. $v_j$ is a random variable assumed to be *iid.* $N(0, \sigma_v^2)$ and independent of $u_j$. $u_j$ is a non-negative random variable, associated with technical inefficiency of production and assumed to be independently distributed as a truncation at zero of the $N(Z_j\delta, \sigma_2)$ distribution. $Z_j$ is a vector of explanatory variables associated with technical inefficiency of production and $\delta$is a vector of unknown coefficients.

The technical inefficiency effect $u_j$ presented in the stochastic frontier model (4) is specified in the following equation:

$$u_j = Z_j\gamma + w_j, \qquad (5)$$

where $w_j$ is defined as the truncation of the $N(0, \sigma_2)$ distribution, such that the point of truncation is $-Z_j\delta$. The predicted technical efficiency of IPU j is given by:

$$TE_j = \frac{E(Y_j|u_j;X_j)}{E(Y_j|u_j=0;X_j)} = \exp(-u_j) \qquad (6)$$

where $X_j$ is the vector containing inputs, year and sector dummies. This transformation bounds the inefficiency term in the [0,1] interval. It implies that a score equal to 1 indicates efficiency or "frontier" technology, and a score less than 1 implies inefficiency of the considered firm.

$Z_j$ is a vector of the following explanatory variables which includes proxies of sharing norms and allocation of time within the household running IPU *j*: the share of the owner's ethnic group in the neighborhood; the log distance to the district of origin, in quadratic form; a dummy indicating whether he/she spent time providing free services to the community in the previous week; the log number of hours spent performing domestic tasks during the previous week; the number of elderly (older than 65 years old) in the household, excluding the owner of the IPU and his/her spouse; the number of children younger than five in the household.[25]

Other controls include the human capital of the entrepreneur and age of the enterprise to capture the managerial ability of the owner. Year dummies capture potential time-varying inefficiency effects. We hypothesize that household characteristics may impact efficiency but have no direct effect on the level of output of the firm, an assumption that we tested and verified in the production function estimation. Therefore, we also add controls for household characteristics by including the factors defined by the PCA, described in Section 2. All specifications include factors 1, 3, 4 and 5, controlling thereby the household size, the migratory status, ethnic group and multiple business ownership of the household. Factor 2, correlated with wealth is not included because of its potential endogeneity. Equations (4) and (5) are estimated simultaneously by maximum likelihood using the computer program NLOGIT 5 (Coelli, 1994).

In a first step, we run separate regressions on male and female entrepreneurs.[26] This specification also includes sector indicators, the share of family labor in total labor and a variable indicating if the activity is located in the home of the entrepreneur.

Correcting for endogenous selection into informal entrepreneurship in stochastic frontier analysis with efficiency determinants is challenging. We follow Greene (2010) who has developed a model that consistently incorporates sample selection into a stochastic frontier model by extending the Heckman selection model to this special non-linear case. However, this model does not allow for simultaneous estimation of inefficiency effects. We therefore estimate inefficiency effects in a second step using OLS with the predicted technical efficiency obtained in the first stage. As this poses a consistency problem[27], we also estimate the Battese and Coelli (1995, thereafter BC) model to which we added the Inverse Mills Ratio as an addition regressor in the stochastic frontier model. Neither of the two options are completely satisfying but by running both and comparing them to the non-corrected model, we increase our confidence that our estimates are plausible (see Wollni and Brümmer, 2012 for a similar approach).

In a second step, we refine these regressions taking into account specific categories of entrepreneurs. We first split the sample into entrepreneurs who are home based and those who have an outside location for their activity, as this can be an important explanatory variable for firm efficiency. Another circumstance under which these variables may have a differentiated impact is running several firms in the same household. Running several firms may be a household strategy to diversify risks. It is also likely that running several firms occurs when households split their activities as a strategy to avoid abusive demands from the extended family, because it is easier to hide several smaller enterprises than one large enterprise. Then, household managing two or more informal firms may be more likely to endure redistributive pressure so that the determinants of technical efficiency may differ according to these two types of households.

Other dimensions are also investigated. We split the sample according to the sector of activity and the sex, the type of IPU (defined by its labor organization: self-employment, non-wage or wage[28]) and the size (defined by the tercile of value added to which the IPU belongs).

Table 10 shows that our estimates are robust when selectivity is accounted for. In the BC model the results are very similar to the non-corrected BC model, while using the Greene method changes the significance of the coefficients but not their sign. For the sub-sample models we refrain from correcting from sample selection. In addition, convergence is difficult to obtain in these selection correction models, all the more when sample sizes are quite small. As this would be too much to ask of the data, we do not run these selectivity corrected models.

We provide a simple test of the relevance of the stochastic frontier analysis in our dataset. The Phase 2 questionnaire includes a question on the possibility for the entrepreneur to increase his/her production without changing the level of capital or labor. Owners who answer yes to this question then declare the potential percentage increase in production, providing a direct measure of the effective under-utilization of the IPU's productive capacity. For such entrepreneurs (about 40 percent of the sample), we ran a bivariate regression of the predicted efficiency using stochastic frontier analysis against the potential percentage increase in production. We obtained a negative and

significant coefficient (at the 5 percent level) confirming that the predicted efficiency indeed corresponds to a sub-optimal use of inputs: efficiency increases as the (perceived) potential increase in production decreases. The stochastic frontier analysis will provide estimates of efficiency for all entrepreneurs, including those who do not perceive that they are using their inputs in a technically inefficient way.[29]

Before we turn to the econometric results, let us first briefly describe the efficiency score predicted by the stochastic frontier regression along three dimensions: gender, location of the activity and ownership of several firms (Table 9). With a mean predicted efficiency of 0.43, female entrepreneurs are significantly less efficient than their male counterparts (0.48) (see Table 9). These scores are in the same order of magnitude as what is reported for Ghanaian manufacturing firms (Söderbom and Teal, 2004) or fish farms (Onumah et al., 2010). Lower female efficiency is not true in every sub-sample however. Home-based females are significantly more efficient than home-based males, and the reverse is true if the activity is located outside the home. Along the multiple IPU ownership dimension, females entrepreneurs are always less efficient than males. Consistent with our assumption, both female- and male-owned businesses belonging to multiple IPU households are less efficient that those in single IPU households.

### (b) Determinants of firm inefficiency across gender

The two blocks of explanatory variables, proxying sharing norms and allocation of time within the household, will be now presented and commented sequentially. The dependent variable is inefficiency, therefore, in Tables 10 to 13, a negative sign should be interpreted as an efficiency-improving effect.

#### (i) Effect of proxies of sharing norms

In the first model, we see, for males only, significant effects of the share of the owner's ethnic group in his neighborhood and of the distance to his district of origin (Table 10, Columns 1 and 4). When we split the sample according to the location of the activity, the share of the ethnic group significantly increases the inefficiency of female entrepreneurs when they are home-based (Table 11). When she works in her home, members of her community, neighbors or friends could take advantage of a social call to ask the owner for favors or help for example. In an outside location, such as a market, pressures are made more difficult to exert because demands would have to be made in public rather than in the privacy of a home. Another possible interpretation of this result is that the entrepreneur is simply distracted by the physical presence of her community and has to fulfill social commitments which have a negative effect on her efficiency.[30]

Only females owning textile and service IPUs, and males running transport businesses are significantly affected by this local network proxy (Table 12). Such activities may be more subject to sharing norms because of the type of good or service produced. A female tailor or hairdresser may feel obliged to sew dresses for her friends and family, or do their haircuts and hairdos. The taxi driver (transport sector) could take time to drive relatives around for free rather than take a paying client.[31] In other sectors, like trade or catering, demands may be less easy to make because

the premises of the business allow less privacy. Besides, the goods or services produced in these other sectors may be less subject to sharing because they do not require skills as specific as haircutting, for example, or special equipment, such as an automobile. When we split the sample according to the tercile of value added, we see that the effect is significant for female owners only if their business is in the first tercile of valued added, suggesting stronger sharing norms in poorer communities (Table 13, Column 2).

The distance variable has a significant quadratic effect for male entrepreneurs (Table 10). Beyond the turning point, the further away they are from their district of origin, the more efficient. The turning point is 23 minutes, which roughly corresponds to the minimum distance for migrants. Below that point, the variable captures the effect of not being a migrant. We tested the model on the sub-sample of migrants (791 observations) and the effect proved very robust. As mentioned in Section 4.1, the migratory status of the household is controlled for by the inclusion of the fourth factor obtained using principal component analysis, which is strongly correlated to the number of migrants in the household and the migratory status of the household head. The positive effect of distance on efficiency is consistent with the assumption that a longer distance makes it more difficult and costly to observe the entrepreneur's activities and productivity and hence redistributive pressure should decline with distance. The effect holds in multiple IPU households but not in households with only one informal business. In fact, in single IPU households, we do not see any effect of the redistributive pressure variables. According to our initial hypothesis, households subject to redistributive pressure are likely to split their activities into several businesses to make their profits less visible to the community. Our results suggest that perhaps this strategy is not effective or that the exerted pressure to share is so strong that this strategy can only mitigate the effect but not completely eliminate it. In this case, the estimated coefficients would be a lower bound of the potential effect of sharing norms on such households.

Similarly to females regarding the local network, the distant network has an effect on males running IPUs in the first tercile of value added (Table 13, Column 1). Again, if poorer communities have stronger sharing norms, this could explain the effect, although for males it happens mainly through the distant rather than the local network. Furthermore, this effect is significant for males who run non-wage IPUs, that is, businesses in which workers are mostly family members. As members of the network who live closer by could have better knowledge of the enterprise and thus know that family members already work in the IPU, they could exert pressure on the owner to hire other relatives and thus decrease the efficiency. They are also more likely to move to the city to work in the business than if they lived in very distant districts. Such effects may push the owner to use more labor than necessary, which reduces his/her efficiency.

Finally, we see an opposite effect of the distance variables for males in the service sector, indicating that in such activities, where competition is strong, a close-by network may be necessary to insure a clientele for example (Table 12, Column 8). The positive effect of social networks has been shown in the literature, and this result confirms that both adverse and beneficial effects have to be considered (see eg. Fafchamps and Minten, 2002). In fact,

in the specifications where such variables are not significant, negative effects of the network could be cancelling out its positive impact.

In all the specifications, only male entrepreneurs appear to feel an impact of the distant network on their efficiency, an explanation of which we attempt to find by looking at the respective characteristics of male and female migrants in the dataset. Firstly, almost half of male migrations were motivated by the prospect of finding a job in the city.[32] The literature on the determinants of migration and transfers argues that migration can be part of household's collective strategy to diversify income sources and remittances made by the migrant would be part of a contractual arrangement between him and the family (see e.g. Lucas and Stark, 1985). Female migration, on the other hand, is mainly motivated by following or joining one's family (60 percent of female migrants). Expectations from the family in the location of origin could then be lower as the migrant did not relocate to increase her earnings. In addition, if she joined her family which had migrated previously, the remaining members of the kin group in the location of origin may be less numerous, if there are any left. Secondly, wives of migrants are more often migrants themselves (72 percent) than the reverse (66 percent of female migrant's husbands). The fact that, for female owners, often both spouses are migrants may mitigate the effect of the pressure to redistribute to the family in the district of origin. Finally, the average distance to the district of origin is higher for female entrepreneurs than for their male counterparts. It is possible that a desire to emancipate themselves from the prevailing gender norms is pushing women to move further away from their relatives. We would then be in the presence of a self-selection effect, as women who wish to free themselves from traditional norms and values are also less likely to conform to other types of cultural norms, such as sharing with the family or the kin group.

Turning to the variable indicating whether the owner provided free services to the community in the previous week, we notice it has a significant negative effect on the efficiency of female entrepreneurs living in multiple IPU households (Table 11, Column 6). One possible explanation is that they are allocating part of their most efficient time to these services, and therefore are less productive in their own activity. Devoting time to the community can be considered a direct effect of solidarity norms as it reduces the available efficient time for production, while the share of the ethnic group or the distance to the district of origin are proxies of more indirect effects of the community on technical efficiency. Noteworthy is the fact that the effect of services for the community is also true in the textile sector (females), transport sector (males) and in non-wage IPUs (males), these three categories exhibiting also significant effects of either the local or the distant network. This suggests an environment in which sharing norms are strong and unfavorable to the entrepreneur. The fact that the direct solidarity variable is significant when the indirect ones are comforts the assumption that the indirect transmission channels are somehow linked to services given to the community.

In the textile sector, the effect of male entrepreneurs providing free services for the community has a reverse effect, as it increases efficiency (Table 12, Column 3). We would interpret this result as the possible existence of reciprocity when aid has been provided to the community.

(ii) Effect of time allocation within the household

We now turn to the second block of variables, which pertains to the role of the household in explaining the efficiency of an entrepreneur. Perhaps surprisingly, the number of hours spent performing domestic tasks has no effect on the efficiency of either sexes (Table 10). However, splitting the sample into home-based and not home-based entrepreneurs again unfolds effects specific to the gender and to the location of the activity. We find that domestic tasks have a negative effect only for females working from home (Table 11, Column 2). The burden of domestic tasks interferes with the productive activity of women when they are at home as they do not have the possibility to disconnect their multiple activities, domestic and market-oriented. Full sample regressions had showed a significant negative effect of having the business located at home for males only (Table 10), a result also found in the descriptive statistics showed in Table 9. Home based activities are therefore not a direct source of inefficiency for female entrepreneurs, but are a vector of transmission of negative externalities on the business management. The negative effect of domestic tasks is particularly true of females running textile businesses, which is essentially done at home, and mostly consists of sewing and embroidering (Table 12, Column 4). While performing these activities, women may feel it is more difficult to compartmentalize their time. Solidarity effects were found to matter in this industry in the previous section, suggesting an unfavorable environment for efficiency, as the entrepreneur is easily distracted. Perhaps the productive process in itself, being quite traditional, can be more easily interrupted, but these frequent interruptions result in a less intense activity.

Domestic tasks negatively impact the efficiency of male entrepreneurs only if they own non-wage IPUs. It is also in this category that males perform the highest average number of hours of domestic tasks. The presence of relatives working in their business, to whom they can delegate supervisory tasks, enables them to perform more domestic chores than the pure self-employed or heads of wage IPUs (with salaried workers). But family workers are often the wife or the children of the owner, who could simply be less productive in market-oriented activities than men.[33,34]

None of the household composition variables have a significant effect in the full model. The presence of elderly has a negative effect on male entrepreneurs in households in which there is just one informal business (Table 11) and in female-run non-wage IPUs (Table 13). The presence of an elderly in the home may entail a need to take care of this person. Another possible explanation is the possible interference of the elderly in the activity of the entrepreneur, suggesting or imposing other, perhaps more traditional, work methods. Owning several IPUs could enable the entrepreneurs to mitigate the negative effects of the elderly on efficiency. In addition, as non-wage IPUs are family businesses, older family members may feel they have more legitimacy in interfering with the entrepreneur's activity. In addition, they can be part of the enterprise's labor force but as they are older, their productivity is lower. Another explanation can be found in the adverse incentive effect of sharing norms within the household (rather than within the local network) as business owners may reduce their productive effort when facing the prospect of sharing their profits with a parent.

Lastly, the number of children in the household that are younger than five years old increases the efficiency of females in single IPU households (Table 11, Column 8). This unexpected effect may seem counter-intuitive, but there are a number of mechanisms that could be at play to explain it. Firstly, it could reflect an unobserved ability bias, by which only the most able women choose to have both several young children and an informal activity. Second, the positive effect of children on efficiency could capture a wealth effect of the household. If there are more than two children under five years old in a household, it is likely that some are fostered children in wealthier households. Such households may also be larger, with several females of childbearing age, who can take care of each other's children. Finally, women who have several young children could be selecting themselves into activities in which efficiency is more easily attainable, making them *de facto* closer to their production frontier. These would be activities with little technology and capital stock for example, allowing them to combine a productive activity with childbearing and childcare. We see in Table 13 that the effect holds in the second tercile but not in the third, which would be consistent with this last explanation, as female self-select into IPUs with less value added (with potentially less technology and capital), while in richer businesses, where value added is higher, children can become a obstacle to efficiency. Finally, we see for males in food and other industries a negative effect of the number of children on their efficiency. Self-consumption of part of the production (food for the children for example) could account for that effect.

## V. DISCUSSION AND CONCLUSION

In this paper, we use a representative sample of informal businesses in Antananarivo to add new evidence on the magnitude of the gender performance gap for informal entrepreneurs in the case of Madagascar. After having highlighted the extent of the adjusted gender performance gap, we suggest possible explanations for the persistence of an unexplained part of the gap. We examine the impact of solidarity norms and gender-differentiated allocation of time within the household on female and male entrepreneurs.

The estimated gender performance gap, after controlling for factor inputs endowment, sectors and the owner's human capital, is 28 percent. Correcting for endogenous selection into informal self-employment raises the gap by 5 percentage points. Returns to capital in female-owned firms are significantly lower than in male-owned firms, while her hours of labor are more productive than his. Quantile regressions confirm that the returns to capital are lower for females than for males all along the distribution of value added. By contrast, returns to labor of the owner are decreasing along the distribution for both women and men and these returns are always higher for females.

We then estimate the extent to which the gender performance gap is due to differences in factor endowments, human capital and sectors, versus differences in the returns to these attributes. A quarter of the gap remains unexplained by differences in mean observable characteristics. When selectivity is accounted for, the unexplained share of the gap drops to 15 percent. The share of the gap that is unexplained increases as we move up the distribution

of value added, suggesting the presence of a glass ceiling phenomenon among female self-employed informal workers, which is not the case among formal wage workers, as shown in Nordman and Wolff (2009b).

In the second empirical part of the paper, we investigate the extent to which household- and community-level gender norms can explain the remaining share of the gap. To this end, we use stochastic frontier analysis to estimate the determinants of technical inefficiency and examine two sets of variables related to sharing norms and time-allocation within the household.

We find evidence of gender-specific effects of sharing norms and within-household time-allocation variables on technical efficiency. Only male entrepreneurs seem subject to pressure to redistribute from the distant network, which could be due to differences in the type of migration of females and males, as women move further away and more often for family, rather than economic reasons. This hints at the presence of unobservable characteristics among female migrants who could be severing their kinship ties to emancipate themselves from prevailing gender norms and would also be less likely to conform to other types of cultural norms, such as sharing with the family or the kin group. Second, while for females, having the business located at home is not a handicap per se, operating from home acts as a vector of transmission of negative externalities due to intense social norms and domestic obligations on the business management. Third, young children in the household have an unexpected positive effect on the efficiency of females running specific categories of businesses. Women could be selecting themselves into industries in which efficiency is more easily attainable to combine their productive and domestic activity, in particular childbearing.

Evidence is found that certain sectors of activity are more subject to redistributive pressure than others, particularly that originating from the local network. The type of goods and services produced in these activities, textile, services and transport are more liable to demands from the close-by community, because of certain specific skills or equipment required to produce them. Both sharing norms and household variables appear to impact non-wage IPUs, which are mainly family businesses, while IPUs with paid workers do not seem affected by any of the proxies used. Non-wage IPU enterprises are more embedded in the household and seem more subject to inefficiencies related to the social network and the composition of the household.

Lastly, we find evidence, in certain segments of the informal sector, for male entrepreneurs, of solidarity norms also acting as beneficial mechanisms.

To substantiate these findings, further research is needed in various directions. One could look at the effect of gender differentiated allocation of expenses within the household which may cause lower profit reinvestment for women, if they tend to take on a bigger share of household expenses such as education and health. Another interesting path to follow would be to investigate whether the performance of male-owned firms has an impact on the level of output and inputs of their wives' businesses.

REFERENCES


Ackerberg, D., Lanier Benkard, C., Berry, S. and Pakes, A. (2007) Econometric Tools for Analyzing Market Outcomes, in J.J. Heckman and E.E. Leamer (eds.), *Handbook of econometrics*, volume 6A, 4171-4276, Elsevier.

Aguirregabiria, V. (2009) Econometric Issues and Methods in the Estimation of Production Functions, MPRA Paper 15973, University Library of Munich, Germany.

Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977) Formulation and Estimation of Stochastic Frontier Production Models. *Journal of Econometrics*, 6, 21-37.

Albrecht, J., Bjorklund, A. and Vroman, S. (2003) Is There a Glass Ceiling in Sweden? *Journal of Labor Economics*, 21(1), 145-177.

Aterido, R., Beck, T. and Iacovone, L. (2011) Gender and Finance in Sub-Saharan Africa: are Women Disadvantaged? Policy Research Working Paper Series 5571, World Bank.

Aterido, R. and Hallward-Driemeier, M (2011) Whose Business is it Anyway? *Small Business Economics*, 37(4), 443-464.

Bardasi, E., Sabarwal, S. and Terrell, K. (2011) How do Female Entrepreneurs Perform? Evidence from Three Developing Regions. *Small Business Economics*, 37(4), 417-441.

Battese, G.E. (1997) A Note on the Estimation of Cobb-Douglas Production Functions when Some Explanatory Variables have Zero Values. *Journal of Agricultural Economics*, 48(1-3), 250-252.

Battese, G.E. and Coelli, T.J. (1995) A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics*, 20, 325-332.

Becker, G. (1981) *A Treatise on the Family*, Harvard University Press.

Bollinger, C., Ziliak J.P. and Troske K.R. (2011) Down from the Mountain: Skill Upgrading and Wages in Appalachia, *Journal of Labor Economics*, 29(4), 819-857.

Buchinsky, M. (2001) Quantile regression with sample selection: Estimating women's return to education in the U.S., *Empirical Economics,* 26, 87–113.

Bourguignon, F., Fournier, M., and Gurgand, M. (2007) Selection Bias Corrections Based on the Multinomial Logit Model: Monte-Carlo Comparisons, *Journal of Economic Surveys*, 21(1), 174–205.

Coelli, T.J. (1994) A Guide to FRONTIER Version 4.1: a Computer Program for Stochastic Frontier Production and Cost Function Estimation. Mimeo, Department of Econometrics, University of New England, Armidale.

Comola, M. and Fafchamps, M. (2010) Are Gifts and Loans Between Households Voluntary?, CSAE Working Paper Series 2010-20.

De La Rica, S., Dolado, J.J. and Llorens, V. (2008) Ceilings or Floors? Gender Wage Gaps by Education in Spain. *Journal of Population Economics*, 21(3), 751-776.

De Mel, S., Mckenzie, D. and Woodruff, C. (2008) Returns to Capital in Microenterprises: Evidence from a Field Experiment. *Quarterly Journal of Economics*, 123(4), 1329-1372

De Mel, S., Mckenzie, D. and Woodruff, C. (2009) Are Women more Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns. *American Economic Journal: Applied Economics*, 1(3), 1-32.

Deolalikar, A.B. and Vijverberg, W.P.M. (1987) A Test of Heterogeneity of Family and Hired Labour in Asian Agriculture. *Oxford Bulletin of Economics and Statistics*, 49(3), 291-305.

Duflo, E. (2003) Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *The World Bank Economic Review*, 17(1), 1-25.

Duflo, E. and Udry, C. (2004) Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices. NBER Working Papers 10498.

Fafchamps, M., Mckenzie, D., Quinn, S. and Woodruff, C. (2011) Female Microenterprises and the Fly-paper Effect: Evidence from a Randomized Experiment in Ghana. CSAE Discussion Paper.

Fafchamps, M. and Minten, B. (2002) Returns to Social Network Capital Among Traders. *Oxford Economic Papers*, 54(2), 173-206.

Grasmuck, S. and Espinal, R. (2000) Market Success or Female Autonomy? *Gender & Society*,14(2), 231-255.

Greene, W. (2010) A stochastic frontier model with correction for sample selection. *Journal of Productivity Analysis*, 34, 15–24.

Grimm, M., Gubert, F., Koriko, O., Lay, J. and Nordman, C.J. (2013) Kinship-ties and Entrepreneurship in Western Africa, *Journal of Small Business and Entrepreneurship*, 26, 125-150.

Grimm, M., Krüger, J. and Lay, J. (2011) Barriers to Entry and Returns to Capital in Informal Activities: Evidence from Sub-Saharan Africa. *Review of Income and Wealth*, 57, S27-S53.

Hoff, K. and Sen, A. (2006) The Kin as a Poverty Trap. *In:* S. Bowles, S. Durlauf and K. Hoff (eds.), *Poverty Traps*, 95-115, Princeton University Press, New York.

Jellal, M., Nordman, C.J. and Wolff, F.C. (2008) Evidence on the Glass Ceiling Effect in France using Matched Worker-firm Data. *Applied Economics*, 40(24), 3233-3250.

Johnston, D. and Le Roux, H. (2007) Leaving the Household out of Family Labour? The Implications for the Size-Efficiency Debate. *The European Journal of Development Research*, 19(3), 355-371.

Jütersonke, O. and Kartas, M. (2010) *Peace and Conflict Impact Assessment (PCIA) for Madagascar,* Centre on Conflict, Development and Peacebuilding, The Graduate Institute, Geneva.

Kevane, M. and Wydick, B. (2001) Microenterprise Lending to Female Entrepreneurs: Sacrificing Economic Growth for Poverty Alleviation? *World Development*, 29(7), 1225-1236.

Koopmans, T.C. (1951) An Analysis of Production as an Efficient Combination of Activities. *In:* Koopmans, T.C. (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics Monograph No. 13., New York.

La Ferrara, E. (2007) Family and Kinship ties in Development: An Economist's Perspective. Paper presented at the 5th AFD-EUDN Conference, Paris.

Lee, L.F. (1983) Generalized Econometric Models with Selectivity, *Econometrica*, 51(2), 507–512.

Lucas, R. and Stark, O. (1985) Motivations to Remit: Evidence from Botswana. *The Journal of Political Economy*, 901-918.

Luke, N. and Munshi, K. (2006) New Roles for Marriage in Urban Africa: Kinship Networks and the Labor Market in Kenya. *The Review of Economics and Statistics,* 88(2), 264-282.

Machado, J.A.F. and Mata, J. (2005) Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression. *Journal of Applied Econometrics*, 20(4), 445-465.

Mahieu, F.R. (1990), *Les fondements de la crise économique en Afrique*, L'Harmattan, Paris.

Mead, D.C. and Liedholm, C. (1998) The Dynamics of Micro and Small Enterprises in Developing Countries. *World Development,* 26(1), 61-74.

Meeusen, W. and Van Den Broeck, J. (1977) Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, 18, 435-444.

Neuman, S. and Oaxaca R. (2004) Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note, *The Journal of Economic Inequality*, 2(1), 3-10.

Neumark, D. (1988) Employers' Discriminatory Behavior and the Estimation of Wage Discrimination. *Journal of Human Resources*, 23(3), 279-295.

Nguyen, C.H. and Nordman, C.J. (2012) Entrepreneurship and Social Networks in Vietnam: Panel Data Evidence. Paper presented at the 7th IZA/World Bank Conference: Employment and Development, New Delhi, India.

Nichter, S. and Goldmark, L. (2009) Small Firm Growth in Developing Countries. *World Development*, 37(9), 1453-1464.

Nordman, C.J., Rakotomanana, F. and Robilliard, A.S. (2010) Gender Disparities in the Malagasy Labor Market. I*n:* J.S. Arbache, A. Kolev and E. Filipiak (eds.), *Gender Disparities in Africa's Labor Market*, 87-154, AFD and World Bank, Africa Development Forum.

Nordman, C.J., Robilliard, A. and Roubaud, F. (2011) Gender and Ethnic Earnings Gaps in Seven West African Cities. *Labour Economics*, 18, S132-S145.

Nordman, C.J. and Roubaud, F. (2009) Reassessing the Gender Wage Gap in Madagascar: Does Labor Force Attachment Really Matter? *Economic Development and Cultural Change*, 57(4), 785-808.

Nordman, C.J. and Wolff, F.C. (2009a) Is there a Glass Ceiling in Morocco? Evidence from Matched Worker-Firm Data. *Journal of African Economies*, 18(4), 592-633.

Nordman, C.J. and Wolff, F.C. (2009b) Islands Through the Glass Ceiling? Evidence of Gender Wage Gaps in Madagascar and Mauritius. *In:* R. Kanbur and J. Svejnar (eds.), *Labor Markets and Economic Development*, 521-544, Routledge Studies in Development Economics.

Onumah, E.E., Bruemmer, B. and Hoerstgen-Schwark, G. (2010) Productivity of Hired and Family Labour and Determinants of Technical Inefficiency in Ghana's Fish Farms. *Agricultural Economics-Czech*, 56, 79-88.

Pasquier-Doumer, L. (2013) Intergenerational Transmission of Self-Employed Status in the Informal Sector: A Constrained Choice or Better Income Prospects? Evidence from Seven West African Countries. *Journal of African Economies,* 22(1), 73-111.

Platteau, J.P. (2000) *Institutions, Social Norms and Economic Development*, Harwood Academic Publishers, Amsterdam.

Rakotomanana, F., Ramilison, E. and Roubaud, F. (2003) The Creation of an Annual Labour Force Survey in Madagascar. An Example for Sub-Saharan Africa. *InterStat*, 27, 35-58.

Ramamonjisoa, J. (2002) Les relations entre « ethnies » à Madagascar: une problématique souvent mal posée. *Afrique Contemporaine*, 202-203, 55-71.

Randrianja, S. and Ellis, S. (2009) *Madagascar: a Short History*, The University of Chicago Press.

Razafindrakoto, M., Roubaud, F. and Torelli, C. (2009) Measuring the Informal Sector and Informal Employment: the Experience Drawn from 1-2-3 Surveys in African Countries. *African Statistical Journal*, 9(Special Issue), 88-147.

Razafindratsima, N. (2005) *Les solidarités privées dans l'agglomération d'Antananarivo (Madagascar) en 1997 :* famangiana *(visites), cohabitation, entraide financière et matérielle.* Ph.D. thesis, Institut d'Etudes Politiques de Paris.

Roubaud, F. (2000) *Identités et transition démocratique: l'exception malgache?* Editions L'Harmattan.

Söderbom, M. and Teal, F. (2004) Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data. *Journal of Development Economics*, 73(1), 369-394.

Vaillant, J., Grimm, M., Lay, J. and Roubaud, F. (2011) Informal Sector Dynamics in Times of Fragile Growth: the Case of Madagascar. Document de Travail DIAL 2011-10.

Wachsberger, J.M. (2009) *L'intégration sociale hiérarchisée. L'exemple d'une métropole en développement: Antananarivo.* Ph.D. thesis, Ecole des Hautes Etudes en Sciences Sociales.

Whitehouse, B. (2011) Enterprising Strangers: Social Capital and Social Liability Among African Migrant Traders. *International Journal of Social Inquiry,* 4(1), 93-111.

Wollni, M. and Brümmer, B. (2012) Productive efficiency of specialty and conventional coffee farmers in Costa Rica: Accounting for technological heterogeneity and self-selection. *Food Policy,* 37, 67–76

Ypeij, A. (2000) *Producing Against Poverty: Female and Male Micro-Entrepreneurs in Lima, Peru*, Amsterdam University Press.

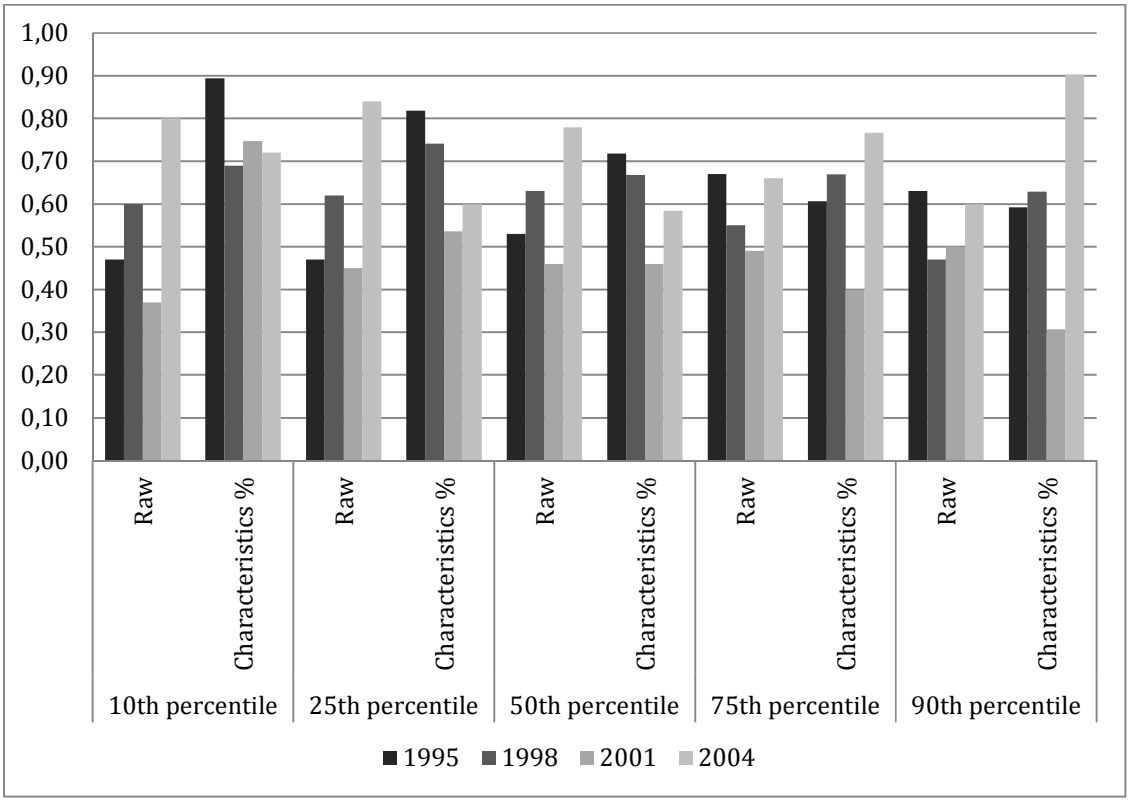


FIGURE 1: RAW GAP AND SHARE OF GAP EXPLAINED BY OBSERVABLE CHARACTERISTICS BY YEAR

Source: *1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO.

TABLE 1: MEAN CHARACTERISTICS OF IPUs AND THEIR OWNERS, BY SEX

| | Male-owned | Female-owned | Total |
|---|---|---|---|
| **IPU characteristics** | | | |
| Value added (log) | 5.70 | 4.98 | 5.34 |
| Age of the enterprise | 8.9 | 8.6 | 8.7 |
| Size of firm (total staff) | 1.6 | 1.35 | 1.47 |
| Pure self-employment (d) | 0.64 | 0.778 | 0.707 |
| Family busness (d) | 0.23 | 0.151 | 0.188 |
| Activity is located in the home | 0.17 | 0.362 | 0.267 |
| Owner's partner works in firm | 0.138 | 0.051 | 0.095 |
| Owner's partner works in firm (if married) | 0.166 | 0.073 | 0.124 |
| Share of family labor (% hours) | 0.142 | 0.084 | 0.113 |
| Firm is the secondary activity of the owner | 0.119 | 0.096 | 0.108 |
| **Inputs** | | | |
| Capital (1000 MGF) | 3685.351 | 1584.554 | 2627.093 |
| Total labor (monthly hours) | 293.289 | 215.471 | 254.089 |
| Owner's labor (monthly hours) | 184.052 | 162.764 | 173.328 |
| Family labor (monthly hours) | 72.499 | 37.449 | 54.843 |
| Hired labor (monthly hours) | 36.739 | 15.259 | 25.918 |
| **Sector** | | | |
| Food processing | 0.025 | 0.025 | 0.025 |
| Textile & clothing | 0.057 | 0.263 | 0.161 |
| Other industry | 0.136 | 0.028 | 0.082 |
| Construction | 0.137 | 0.002 | 0.069 |
| Primary goods trade | 0.128 | 0.244 | 0.186 |
| Transformed good trade | 0.166 | 0.168 | 0.167 |
| Services | 0.231 | 0.209 | 0.220 |
| Catering | 0.021 | 0.058 | 0.040 |
| Transport | 0.099 | 0.002 | 0.050 |
| **Owner characteristics** | | | |
| Age of owner | 0.384 | 0.384 | 0.384 |
| Completed primary school | 0.307 | 0.365 | 0.336 |
| Completed middle school | 0.217 | 0.199 | 0.208 |
| Completed high school | 0.166 | 0.096 | 0.131 |
| Experience | 10.44 | 9.05 | 9.740 |
| Owner married (d) | 0.818 | 0.674 | 0.745 |
| Migrant (d) | 0.286 | 0.280 | 0.283 |
| **Household characteristcs** | | | |
| Size | 5.507 | 5.422 | 5.464 |
| Dependency ratio (inactive/active) | 1.380 | 1.240 | 1.309 |
| Another member in public sector (d) | 0.058 | 0.135 | 0.097 |
| Another member in private formal sector (d) | 0.234 | 0.330 | 0.282 |
| Another member head of IPU (d) | 0.463 | 0.408 | 0.435 |
| Another member employed in IPU (f) | 0.103 | 0.154 | 0.129 |
| Household owns more than one IPU (d) | 0.513 | 0.499 | 0.506 |
| **Proxies of sharing norms and time-allocation within the household** | | | |
| Free services for the community (Hours spent) | 0.116 | 0.085 | 0.100 |
| Distance to district of origin (if migrants, minutes) | 196.893 | 218.191 | 207.511 |
| Share of same ethnic group in neighborhood | 0.856 | 0.824 | 0.840 |
| Domestic tasks (Hours spent) | 3.74 | 13.64 | 8.730 |
| Domestic tasks, market & wood (hours spent) | 6.131 | 17.686 | 11.856 |
| Number of elderly (> 64 years old) | 0.053 | 0.080 | 0.067 |
| Number of children (< 5 years old) | 0.656 | 0.573 | 0.614 |
| Observations | 2159 | 1733 | 3882 |

*Notes*: Means and shares are calculated using Phase 2 sampling weights. Number of elderly excludes the IPU owner and his/her spouse. (d) indicates a dummy variable. *Source: 1-2-3 Surveys*, Phases 1 and 2, 1995-2004,INSTAT/DIAL/MADIO; author's calculation.

TABLE 2: PRODUCTIONS FUNCTIONS. DEPENDENT VARIABLE: LOG VALUE ADDED

| | (1) OLS | (2) OLS | (3) OLS | (4) Fixed Effects | (5) OLS | | (6) LEE |
|---|---|---|---|---|---|---|---|
| | | | | | X | X*female | |
| Female owner | -0.339*** | -0.286*** | -0.193** | -0.224** | 0.457 | | -0.334*** |
| | (0.043) | (0.043) | (0.085) | (0.109) | (0.548) | | (0.048) |
| **Inputs** | | | | | | | |
| Capital | 0.155*** | 0.127*** | 0.134*** | 0.061** | 0.152*** | -0.053*** | 0.128*** |
| | (0.010) | (0.010) | (0.020) | (0.029) | (0.014) | (0.020) | (0.012) |
| Non-zero capital | -0.741*** | -0.663*** | -0.599*** | -0.322 | -0.872*** | 0.394** | -0.667*** |
| | (0.090) | (0.089) | (0.170) | (0.232) | (0.125) | (0.178) | (0.098) |
| Owner labor | 0.352*** | 0.381*** | 0.341*** | 0.366*** | 0.344*** | 0.081 | 0.380*** |
| | (0.025) | (0.025) | (0.048) | (0.064) | (0.033) | (0.051) | (0.022) |
| Non-zero owner labor | -2.426*** | -2.513*** | -2.570*** | -2.389*** | -1.859*** | -1.193** | -2.510*** |
| | (0.248) | (0.244) | (0.446) | (0.660) | (0.347) | (0.490) | (0.266) |
| Family labor | 0.250*** | 0.258*** | 0.239*** | 0.236** | 0.289*** | -0.062 | 0.261*** |
| | (0.041) | (0.040) | (0.086) | (0.114) | (0.055) | (0.082) | (0.033) |
| Non-zero family labor | -0.833*** | -0.871*** | -0.784* | -0.826 | -1.049*** | 0.361 | -0.879*** |
| | (0.213) | (0.210) | (0.437) | (0.570) | (0.295) | (0.423) | (0.182) |
| Hired labor | 0.321*** | 0.296*** | 0.365*** | 0.324** | 0.339*** | -0.117 | 0.295*** |
| | (0.060) | (0.059) | (0.117) | (0.152) | (0.071) | (0.129) | (0.055) |
| Non-zero hired labor | -0.783** | -0.701** | -0.866 | -1.010 | -0.961** | 0.672 | -0.699** |
| | (0.336) | (0.330) | (0.657) | (0.841) | (0.399) | (0.716) | (0.310) |
| **Human capital of owner** | | | | | | | |
| Age of owner | | 2.169*** | 2.505 | 3.447 | 2.777** | -1.251 | 4.167*** |
| | | (0.803) | (1.616) | (2.188) | (1.135) | (1.626) | (1.339) |
| Age of owner2 | | -0.035*** | -0.040** | -0.051** | -0.043*** | 0.015 | -0.058*** |
| | | (0.009) | (0.019) | (0.025) | (0.013) | (0.019) | (0.016) |
| Primary school | | 0.184*** | 0.161* | 0.035 | 0.149** | 0.052 | 0.148*** |
| | | (0.045) | (0.087) | (0.133) | (0.062) | (0.092) | (0.042) |
| Middle school | | 0.282*** | 0.218** | 0.238 | 0.187*** | 0.200* | 0.218*** |
| | | (0.052) | (0.101) | (0.171) | (0.069) | (0.106) | (0.053) |
| High school | | 0.568*** | 0.543*** | 0.125 | 0.530*** | 0.013 | 0.458*** |
| | | (0.060) | (0.123) | (0.235) | (0.077) | (0.126) | (0.078) |
| Experience | | 0.003 | 0.004 | 0.008 | 0.002 | 0.001 | 0.003 |
| | | (0.002) | (0.005) | (0.007) | (0.003) | (0.005) | (0.002) |
| Married | | 0.051 | -0.059 | 0.023 | 0.100 | -0.093 | 0.096* |
| | | (0.044) | (0.085) | (0.164) | (0.069) | (0.091) | (0.051) |
| Age of IPU | 0.021 | 0.711*** | 0.209 | 0.080 | 0.733** | -0.096 | 0.721** |
| | (0.193) | (0.236) | (0.483) | (0.643) | (0.311) | (0.477) | (0.308) |
| Selection term (Lee) | | | | | | | -0.251** |
| | | | | | | | (0.110) |
| Constant | 5.658*** | 5.098*** | 5.270*** | 5.454*** | 4.706*** | | 4.395*** |
| | (0.234) | (0.273) | (0.506) | (0.711) | (0.390) | | (0.462) |
| Observations | 3,882 | 3,882 | 984 | 984 | 3882 | | 3882 |
| R2 | 0.373 | 0.396 | 0.398 | 0.292 | 0.404 | | |
| F-test: all female interactions = 0 | | | | | 2.001 | | |
| p-value | | | | | 0.002 | | |

*Notes*: Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. All regressions include year and sector indicators. In Columns (3) and (4), regressions are run on the sub-sample of IPUs belonging to households for which observations on several businesses are available. Inputs are expressed in logarithmic form, labor variables are log monthly hours. Age variables are divided by 100 to improve legibility. Education variables are indicators, equal to one if the owner completed the indicated level. The selection term is a generalized form of the Mill's ratio obtained using a multinomial logit model in the first stage of the selection correction model (Lee, 1983). *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 3: PRODUCTION FUNCTIONS BY OWNER'S SEX. OLS AND CORRECTING FOR SELECTION INTO INFORMAL SELF-EMPLOYMENT

| | (1) OLS | (2) LEE | (3) OLS | (4) LEE |
|---|---|---|---|---|
| | Male | Male | Female | Female |
| **Inputs** | | | | |
| Capital | 0.152*** | 0.152*** | 0.101*** | 0.099*** |
| | (0.0138) | (0.010) | (0.0147) | (0.015) |
| Owner labor | 0.344*** | 0.342*** | 0.436*** | 0.424*** |
| | (0.0328) | (0.050) | (0.0386) | (0.032) |
| Family labor | 0.289*** | 0.293*** | 0.229*** | 0.228*** |
| | (0.0544) | (0.049) | (0.0615) | (0.068) |
| Hired labor | 0.339*** | 0.337*** | 0.235** | 0.224 |
| | (0.0695) | (0.067) | (0.108) | (0.138) |
| **Human capital of owner** | | | | |
| Age of owner | 2.777** | 4.717*** | 1.581 | 3.150 |
| | (1.113) | (1.523) | (1.178) | (2.175) |
| Age of owner2 | -0.0425*** | -0.067*** | -0.0267* | -0.046* |
| | (0.0128) | (0.018) | (0.0136) | (0.026) |
| Primary school | 0.149** | 0.089 | 0.214*** | 0.186*** |
| | (0.0608) | (0.074) | (0.0690) | (0.071) |
| Middle school | 0.187*** | 0.100 | 0.400*** | 0.348*** |
| | (0.0680) | (0.076) | (0.0823) | (0.099) |
| High school | 0.530*** | 0.386*** | 0.541*** | 0.473*** |
| | (0.0759) | (0.092) | (0.0996) | (0.144) |
| Experience | 0.00196 | 0.002 | 0.00241 | 0.003 |
| | (0.00318) | (0.003) | (0.00377) | (0.004) |
| Married | 0.100 | 0.231** | 0.0347 | 0.012 |
| | (0.0679) | (0.103) | (0.0598) | (0.067) |
| Age of the enterprise | 0.733** | 0.741** | 0.630* | 0.652** |
| | (0.305) | (0.374) | (0.372) | (0.295) |
| Selection term | | -0.312 | | -0.181 |
| | | (0.196) | | (0.175) |
| Constant | 4.706*** | 3.895*** | 5.197*** | 4.603*** |
| | (0.382) | (0.630) | (0.391) | (0.716) |
| Observations | 21159 | 2159 | 1733 | 1733 |
| R2 | 0.378 | | 0.342 | |

*Notes*: Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. All regressions include year and sector indicators. Age variables are divided by 100 to improve legibility. Education variables are indicators, equal to one if the owner completed the indicated level. The selection term is a generalized form of the Mill's ratio obtained using a multinomial logit model in the first stage of the selection correction model (Lee, 1983). *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 4: MALE ENTREPRENEURS: QUANTILE REGRESSIONS CORRECTED FOR SELECTION INTO INFORMAL SELF-EMPLOYMENT.
DEPENDENT VARIABLE: LOG VALUE ADDED

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | 0.1 | 0.25 | 0.5 | 0.75 | 0.9 |
| **Inputs** | | | | | |
| Capital | 0.111*** | 0.153*** | 0.163*** | 0.181*** | 0.168*** |
| | (0.038) | (0.020) | (0.016) | (0.018) | (0.026) |
| Owner labor | 0.499*** | 0.436*** | 0.363*** | 0.211*** | 0.227*** |
| | (0.082) | (0.038) | (0.047) | (0.043) | (0.057) |
| Family labor | 0.131 | 0.234*** | 0.248*** | 0.438*** | 0.366*** |
| | (0.128) | (0.074) | (0.060) | (0.062) | (0.061) |
| Hired labor | 0.424*** | 0.406*** | 0.404*** | 0.291*** | 0.191** |
| | (0.112) | (0.073) | (0.079) | (0.083) | (0.092) |
| **Human capital of owner** | | | | | |
| Age of owner | 5.527* | 3.897* | 4.262** | 4.837** | 2.304 |
| | (3.073) | (2.016) | (2.052) | (2.354) | (3.947) |
| Age of owner2 | -0.085** | -0.057** | -0.058** | -0.064** | -0.037 |
| | (0.037) | (0.025) | (0.025) | (0.029) | (0.047) |
| Primary school | 0.001 | 0.080 | 0.131* | 0.121 | 0.036 |
| | (0.149) | (0.073) | (0.078) | (0.077) | (0.174) |
| Middle school | 0.015 | 0.082 | 0.130 | 0.148 | 0.103 |
| | (0.143) | (0.101) | (0.109) | (0.094) | (0.174) |
| High school | 0.282 | 0.246** | 0.475*** | 0.417*** | 0.422** |
| | (0.208) | (0.114) | (0.143) | (0.146) | (0.168) |
| Experience | 0.007 | 0.005 | 0.004 | 0.001 | -0.002 |
| | (0.006) | (0.004) | (0.004) | (0.004) | (0.005) |
| Married | 0.058 | 0.140 | 0.218 | 0.389** | 0.478** |
| | (0.137) | (0.089) | (0.171) | (0.156) | (0.199) |
| Age of the enterprise | 0.482 | 0.698* | 0.581 | 0.647 | 0.500 |
| | (0.467) | (0.368) | (0.482) | (0.529) | (0.423) |
| Selection term | -0.158 | -0.182 | -0.260 | -0.509* | -0.320 |
| | (0.301) | (0.185) | (0.300) | (0.307) | (0.367) |
| Constant | 1.803 | 3.624*** | 4.177*** | 4.298*** | 5.855*** |
| | (1.131) | (1.230) | (1.027) | (1.028) | (1.372) |
| Observations | 2159 | 2159 | 2159 | 2159 | 2159 |

*Notes*: Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. All regressions include year and sector indicators. Age variables are divided by 100 to improve legibility. Education variables are indicators, equal to one if the owner completed the indicated level. The selection term is a generalized form of the Mill's ratio obtained using a multinomial logit model in the first stage of the selection correction model (Lee, 1983). *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 5: FEMALE ENTREPRENEURS: QUANTILE REGRESSIONS CORRECTED FOR SELECTION INTO INFORMAL SELF-EMPLOYMENT.
DEPENDENT VARIABLE: LOG VALUE ADDED

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| VARIABLES | 0.1 | 0.25 | 0.5 | 0.75 | 0.9 |
| Capital | 0.085*** | 0.089*** | 0.108*** | 0.109*** | 0.125*** |
| | (0.028) | (0.016) | (0.018) | (0.019) | (0.030) |
| Owner labor | 0.548*** | 0.558*** | 0.426*** | 0.418*** | 0.340*** |
| | (0.080) | (0.034) | (0.052) | (0.040) | (0.065) |
| Family labor | 0.123 | 0.162*** | 0.240*** | 0.203** | 0.200* |
| | (0.136) | (0.054) | (0.045) | (0.083) | (0.105) |
| Hired labor | 0.192 | 0.143 | 0.148 | 0.141 | 0.383** |
| | (0.221) | (0.269) | (0.144) | (0.169) | (0.154) |
| Age of the enterprise | 0.241 | 0.500 | 0.335 | 0.910 | 2.000** |
| | (0.783) | (0.416) | (0.287) | (0.568) | (0.959) |
| Age of owner | 3.839 | 4.576** | 3.183 | 5.894* | -2.069 |
| | (4.410) | (2.149) | (2.516) | (3.170) | (3.186) |
| Age of owner2 | -0.044 | -0.058** | -0.046* | -0.080** | 0.006 |
| | (0.051) | (0.024) | (0.027) | (0.036) | (0.036) |
| Primary school | 0.197 | 0.135** | 0.217** | 0.099 | 0.286** |
| | (0.184) | (0.056) | (0.110) | (0.106) | (0.120) |
| Middle school | 0.356 | 0.220 | 0.330** | 0.246** | 0.539*** |
| | (0.236) | (0.139) | (0.142) | (0.114) | (0.172) |
| High school | 0.408 | 0.407** | 0.478*** | 0.385** | 0.691*** |
| | (0.329) | (0.168) | (0.175) | (0.185) | (0.182) |
| Experience | -0.001 | -0.000 | 0.003 | 0.004 | 0.005 |
| | (0.005) | (0.004) | (0.004) | (0.005) | (0.007) |
| Married | -0.075 | 0.046 | 0.024 | -0.021 | 0.081 |
| | (0.146) | (0.083) | (0.062) | (0.064) | (0.092) |
| Selection term | -0.088 | -0.247 | -0.211 | -0.570* | 0.261 |
| | (0.467) | (0.218) | (0.247) | (0.322) | (0.326) |
| Constant | 3.164*** | 3.172*** | 4.303*** | 4.809*** | 7.709*** |
| | (1.145) | (0.957) | (0.893) | (1.182) | (1.102) |
| Observations | 1733 | 1733 | 1733 | 1733 | 1733 |

*Notes*: Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. All regressions include year and sector indicators. Age variables are divided by 100 to improve legibility. Education variables are indicators, equal to one if the owner completed the indicated level. The selection term is a generalized form of the Mill's ratio obtained using a multinomial logit model in the first stage of the selection correction model (Lee, 1983). *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 6: TECHNICAL RATES OF SUBSTITUTION IN MALE- AND FEMALE-OWNED IPUs

| | | **Elasticities** | **Elasticities from quantile regressions** | | | | |
|---|---|---|---|---|---|---|---|
| | | **from LEE** | 0.1 | 0.25 | 0.5 | 0.75 | 0.9 |
| Capital/Owner labor | Male | 0.018 | 0.015 | 0.031 | 0.027 | 0.033 | 0.013 |
| | Female | 0.017 | 0.018 | 0.018 | 0.027 | 0.024 | 0.012 |
| Owner labor/Family labor | Male | 0.483 | 1.025 | 0.454 | 0.511 | 0.215 | 0.473 |
| | Female | 0.507 | 0.416 | 0.529 | 0.318 | 0.659 | 0.927 |
| Owner labor/Hired labor | Male | 0.302 | 0.056 | 0.044 | 0.086 | 0.226 | 1.214 |
| | Female | 0.316 | 0.042 | 0.063 | 0.206 | 0.424 | 0.509 |
| Family labor/Hired labor | Male | 0.626 | 0.055 | 0.098 | 0.167 | 1.052 | 2.568 |
| | Female | 0.623 | 0.100 | 0.119 | 0.649 | 0.643 | 0.548 |

*Notes:* The table shows the Technical Rates of Substitution calculated using the elasticities estimated by the selectivity corrected regressions. *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 7: NEUMARK AND QUANTILE DECOMPOSITIONS OF THE GENDER PERFORMANCE GAP

| | Not corrected | | Corrected for selection | |
|---|---|---|---|---|
| | Raw diff. | Explained (%) | Raw diff. | Explained (%) |
| | (1) | (2) | (3) | (4) |
| **Neumark decomposition** | | | | |
| *Inputs* | 0.743 | 42.4 | 0.614 | 50.8 |
| *Inputs + IPU* | 0.743 | 69.0 | 0.614 | 77.8 |
| *Inputs + IPU + Owner* | 0.743 | 74.8 | 0.614 | 83.8 |
| **Quantile decomposition** | | | | |
| 10th percentile | 0.761 | 80.8 | 0.639 | 99.7 |
| 25th percentile | 0.748 | 70.7 | 0.614 | 87.3 |
| 50th percentile | 0.763 | 55.7 | 0.637 | 70.2 |
| 75th percentile | 0.740 | 49.6 | 0.608 | 61.1 |
| 90th percentile | 0.693 | 51.4 | 0.556 | 66.7 |

*Notes*: All the decompositions include year dummies. *Inputs, IPU* and *Owner* are the vectors of inputs, IPU characteristics and owner characteristics used in the previous regressions. The quantile decompositions include the full set of observables. *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 8: NEUMARK AND QUANTILE DECOMPOSITIONS OF THE GENDER PERFORMANCE GAP BY YEAR (SELECTIVITY CORRECTED)

| | 1995 | | 1998 | | 2001 | | 2004 | |
|---|---|---|---|---|---|---|---|---|
| | Raw | Expl.(%) | Raw | Expl.(%) | Raw | Expl.(%) | Raw | Expl.(%) |
| **Neumark decomposition** | | | | | | | | |
| *Inputs* | 0.547 | 53.4 | 0.567 | 36.4 | 0.442 | 52.3 | 0.742 | 46.8 |
| *Inputs + IPU* | 0.547 | 82.0 | 0.567 | 71.2 | 0.442 | 75.1 | 0.742 | 70.4 |
| *Inputs + IPU + Owner* | 0.547 | 85.4 | 0.567 | 80.7 | 0.442 | 78.3 | 0.742 | 77.1 |
| **Quantile decomposition** | | | | | | | | |
| 10th percentile | 0.468 | 89.4 | 0.598 | 68.9 | 0.370 | 74.7 | 0.798 | 72.0 |
| 25th percentile | 0.470 | 81.8 | 0.621 | 74.1 | 0.455 | 53.6 | 0.836 | 60.0 |
| 50th percentile | 0.535 | 71.8 | 0.635 | 66.7 | 0.457 | 46.0 | 0.783 | 58.4 |
| 75th percentile | 0.672 | 60.6 | 0.551 | 67.0 | 0.492 | 39.9 | 0.665 | 76.7 |
| 90th percentile | 0.625 | 59.2 | 0.473 | 62.9 | 0.498 | 30.7 | 0.596 | 90.3 |

*Notes*: All the decompositions include year dummies. *Inputs, IPU* and *Owner* are the vectors of inputs, IPU characteristics and owner characteristics used in the previous regressions. The quantile decompositions include the full set of observables. *Source: 1-2-3 Surveys*, Phase 2, 1995-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 9: MEAN PREDICTED EFFICIENCY SCORE, BY SEX

| | Male | Female | Student test of difference (p-value) |
|---|---|---|---|
| All | 0.484 | 0.431 | 0.000 |
| At home | 0.409 | 0.458 | 0.003 |
| Not at home | 0.501 | 0.412 | 0.000 |
| Multiple IPU household | 0.467 | 0.414 | 0.000 |
| Single IPU household | 0.501 | 0.445 | 0.000 |

*Notes*: The table shows the mean technical efficiency scores predicted by the general model presented in Columns (1) and (4) of the next table (Table 10). *Source: 1-2-3 Surveys,* Phase 2, 1998-2004, INSTAT/DIAL/MADIO; author's calculation.

TABLE 10: DETERMINANTS OF FEMALE AND MALE-OWNED IPU INEFFICIENCY WITHOUT AND WITH CONTROLING FOR SELECTION INTO INFORMAL SELF-EMPLOYMENT

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | MALE | | | FEMALE | | |
| INEFFICIENCY EFFECTS | B&C | OLS | B&C | B&C | OLS | B&C |
| SELECTION MODEL | | GREENE | MILLS | | GREENE | MILLS |
| **Sharing norms** | | | | | | |
| Share of same ethnic group | 1.235** | 0.079*** | 1.141** | -0.033 | 0.021 | -0.045 |
| | (0.488) | (0.025) | (0.449) | (0.244) | (0.038) | (0.267) |
| Free services for the community | 0.281 | 0.052* | 0.299 | 0.253 | 0.025 | 0.272 |
| | (0.401) | (0.03) | (0.374) | (0.261) | (0.05) | (0.283) |
| Log distance | 0.402** | 0.024 | 0.389** | -0.168 | -0.034 | -0.186 |
| | (0.204) | (0.015) | (0.189) | (0.158) | (0.025) | (0.174) |
| Log distance squared | -0.065* | -0.004 | -0.062* | 0.017 | 0.004 | 0.019 |
| | (0.038) | (0.003) | (0.035) | (0.028) | (0.004) | (0.031) |
| **Time allocation within the household** | | | | | | |
| Domestic tasks & market | 0.008 | 0.001 | 0.009 | 0.002 | 0.004 | 0.002 |
| | (0.015) | (0.001) | (0.014) | (0.024) | (0.004) | (0.026) |
| Nb of children | -0.026 | -0.002 | -0.023 | -0.032 | -0.013 | -0.035 |
| | (0.109) | (0.008) | (0.102) | (0.081) | (0.014) | (0.088) |
| Number of elderly | 0.098 | -0.024 | 0.041 | 0.534 | 0.07 | 0.546 |
| | (0.318) | (0.031) | (0.299) | (0.424) | (0.053) | (0.428) |
| Activity located at home | 0.593** | 0.042*** | 0.566** | 0.25 | 0.037 | 0.27 |
| | (0.251) | (0.016) | (0.236) | (0.188) | (0.03) | (0.201) |
| Lambda | 1.444*** | | 1.402*** | 0.867*** | | 0.915*** |
| | (0.148) | | (0.143) | (0.131) | | (0.13) |
| Sigma | 1.339*** | | 1.306*** | 1.143*** | | 1.167*** |
| | (0.075) | | (0.068) | (0.041) | | (0.048) |
| N | 1672 | 1672 | 1672 | 1207 | 1207 | 1207 |

*Notes:* Table shows only the results of the inefficiency model, the estimates of the stochastic frontier production function are not presented. Controls not shown but included in the determinants of inefficiency equation: year and sector indicators, human capital of the owner (age, experience, schooling, marital status), age of the IPU, factors 1,3,4 and 5 obtained using the PCA of household characteristics. Columns (1) and (4) show results of the Battese & Coelli (1995) model, Columns (2) and (5) correct for endogenous selection into informal self-employment using the model developed by Greene (2010), Columns (3) and (6) correct for endogenous selection into informal self-employment by adding the Mill's ratio as a regressor (in the frontier) to the Battese & Coelli model.

TABLE 11: DETERMINANTS OF FEMALE- AND MALE-OWNED IPU INEFFICIENCY

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Home based | | Not home based | | Multiple IPU households | | Single IPU households | |
| Variables | Male | Female | Male | Female | Male | Female | Male | Female |
| **Sharing norms** | | | | | | | | |
| Share of same ethnic group | 0.454 | 1.205** | 3.751 | -0.297 | 0.977 | 0.039 | 0.24 | 3.337 |
| | (0.638) | (0.5) | (4.459) | (0.275) | (0.62) | (0.753) | (0.747) | (2.064) |
| Free services for community | -0.133 | 0.48 | 2.274 | 0.102 | 0.295 | 1.81** | 0.233 | -0.533 |
| | (0.372) | (0.356) | (2.435) | (0.236) | (0.433) | (0.858) | (0.454) | (0.927) |
| Log distance | 0.627** | -0.319 | 1.064 | 0.027 | 0.511** | -0.666 | -0.025 | -0.326 |
| | (0.279) | (0.251) | (1.295) | (0.177) | (0.25) | (0.906) | (0.282) | (0.362) |
| Log distance squared | -0.097** | 0.056 | -0.146 | -0.007 | -0.082* | 0.108 | 0.012 | 0.013 |
| | (0.048) | (0.042) | (0.182) | (0.032) | (0.044) | (0.15) | (0.048) | (0.06) |
| **Time allocation within the household** | | | | | | | | |
| Domestic tasks & market | 0.016 | 0.148* | -0.001 | -0.009 | -0.021 | 0.056 | 0.046 | 0.058 |
| | (0.021) | (0.077) | (0.046) | (0.014) | (0.016) | (0.061) | (0.03) | (0.049) |
| Number of children | 0.001 | -0.172 | 0.061 | 0.091 | 0.027 | 0.341 | 0.18 | -2.113** |
| | (0.139) | (0.136) | (0.18) | (0.056) | (0.099) | (0.209) | (0.156) | (1.066) |
| Number of elderly | 0.571 | 0.201 | 0.713 | 0.204 | -0.169 | -0.098 | 1.135** | 1.479** |
| | (0.376) | (0.233) | (0.994) | (0.178) | (0.299) | (0.833) | (0.542) | (0.675) |
| Activity located at home | | | | | 0.619*** | -0.374 | 0.186 | 0.932** |
| | | | | | (0.222) | (0.463) | (0.283) | (0.421) |
| N | 311 | 478 | 1369 | 737 | 846 | 580 | 834 | 635 |

*Notes*: Table shows only the results of the inefficiency model, the estimates of the stochastic frontier production function are not presented. Controls not shown but included in the determinants of inefficiency equation: year indicators, human capital of the owner (age, experience, schooling, marital status), age of the IPU, factors 1,3,4 and 5 obtained using the PCA of household characteristics.

TABLE 12: DETERMINANTS OF FEMALE- AND MALE-OWNED IPU INEFFICIENCY BY SECTOR OF ACTIVITY

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Food & Other ind. | | Textile | | Construction | Trade | | Services | | Catering | Transport |
| | Male | Female | Male | Female | Male | Male | Female | Male | Female | Female | Male |
| **Sharing norms** | | | | | | | | | | | |
| Share of same ethnic group | 0.582 | -0.386 | 0.935 | 1.326** | 1.11 | 3.228 | -0.639 | -0.58 | 1.186* | -1.195 | 1.925** |
| | (0.937) | (0.723) | (0.937) | (0.594) | (0.715) | (3.126) | (0.915) | (0.782) | (0.609) | (1.073) | (0.829) |
| Free services for community | 0.773 | 0.194 | -2.806*** | 0.759* | 0.198 | -0.101 | 0.833 | 0.368 | 0.468 | -2.755 | 1.572* |
| | (0.864) | (0.979) | (1.003) | (0.402) | (0.933) | (2.793) | (1.643) | (0.758) | (0.56) | (1.871) | (0.823) |
| Log distance | 0.256 | -0.482 | 0.573 | -0.356 | -0.494 | 1.261 | -0.476 | 0.862** | 0.244 | 1.227 | -0.292 |
| | (0.874) | (0.548) | (0.737) | (0.241) | (0.315) | (1.139) | (0.909) | (0.404) | (0.371) | (0.842) | (0.305) |
| Log distance squared | -0.073 | 0.008 | -0.064 | 0.053 | 0.075 | -0.207 | 0.054 | -0.114 | -0.03 | -0.236 | 0.054 |
| | (0.159) | (0.085) | (0.125) | (0.041) | (0.054) | (0.189) | (0.13) | (0.075) | (0.06) | (0.149) | (0.047) |
| **Time allocation within the household** | | | | | | | | | | | |
| Domestic tasks & market | 0.009 | 0.022 | -0.018 | 0.122* | 0.012 | -0.044 | -0.124 | 0.002 | 0.059 | -0.072 | 0.007 |
| | (0.033) | (0.047) | (0.055) | (0.068) | (0.022) | (0.064) | (0.115) | (0.034) | (0.061) | (0.109) | (0.03) |
| Number of children | 0.642** | -0.075 | -0.071 | -0.012 | -0.08 | 0.037 | -0.001 | 0.05 | 0.082 | -0.555 | 0.229 |
| | (0.261) | (0.23) | (0.38) | (0.18) | (0.111) | (0.489) | (0.225) | (0.203) | (0.169) | (0.643) | (0.171) |
| Number of elderly | 1.087 | 0.014 | 2.116*** | 0.631 | 0.04 | 1.641 | -0.086 | -0.028 | 0.065 | 2.234* | -0.149 |
| | (0.775) | (0.645) | (0.68) | (0.393) | (0.332) | (1.756) | (0.919) | (0.464) | (0.773) | (1.206) | (0.482) |
| Activity located at home | -0.126 | -0.192 | -1.003** | -0.484 | | 0.479 | 0.441 | 0.808** | 0.196 | | |
| | (0.532) | (0.466) | (0.45) | (0.539) | | (1.091) | (0.786) | (0.338) | (0.349) | | |
| N | 321 | 126 | 93 | 307 | 289 | 301 | 373 | 402 | 250 | 140 | 222 |

*Notes*: Table shows only the results of the inefficiency model, the estimates of the stochastic frontier production function are not presented. Controls not shown but included in the determinants of inefficiency equation: year indicators, human capital of the owner (age, experience, schooling, marital status), age of the IPU, factors 1,3,4 and 5 obtained using the PCA of household characteristics.

TABLE 13: DETERMINANTS OF FEMALE- AND MALE-OWNED IPU INEFFICIENCY BY TERCILES OF VALUE ADDED AND TYPE

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1st tertile | | 2nd tertile | | 3rd tertile | | Self-employment | | Non-wage IPU | | Wage IPU | |
| | Male | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male | Female |
| **Sharing norms** | | | | | | | | | | | | |
| Share of same ethnic group | -0.24 | 1.409** | 0.139 | -0.156 | -0.057 | -0.343 | 5.311* | 0.266 | 0.298 | -0.989 | -0.556 | 1.386 |
| | (0.747) | (0.69) | (0.882) | (0.198) | (0.777) | (0.271) | (3.083) | (0.805) | (0.806) | (1.058) | (0.694) | (0.958) |
| Free services for community | -1.346 | 0.588 | -0.057 | -0.126 | 0.14 | -0.451 | -0.229 | 0.47 | 1.239** | 0.184 | 0.002 | 0.009 |
| | (0.923) | (0.9) | (0.988) | (0.133) | (0.474) | (0.288) | (0.815) | (0.85) | (0.525) | (0.997) | (0.399) | (0.957) |
| Log distance | 1.72*** | 0.034 | -0.083 | 0.079 | 0.052 | 0.038 | 1.485 | 0.094 | 0.582** | 0.011 | 0.098 | -0.957 |
| | (0.434) | (0.776) | (0.767) | (0.07) | (0.48) | (0.138) | (1.035) | (0.607) | (0.26) | (0.722) | (0.223) | (0.787) |
| Log distance squared | -0.323*** | -0.02 | 0.006 | -0.015 | -0.008 | -0.009 | -0.228 | -0.008 | -0.118** | -0.044 | -0.017 | 0.181 |
| | (0.076) | (0.13) | (0.146) | (0.012) | (0.08) | (0.022) | (0.163) | (0.094) | (0.046) | (0.129) | (0.038) | (0.146) |
| **Time allocation within the household** | | | | | | | | | | | | |
| Domestic tasks & market | -0.01 | 0.047 | 0.006 | -0.017** | 0.007 | -0.013 | 0.021 | 0.048 | 0.056* | -0.053 | -0.012 | 0.08 |
| | (0.025) | (0.05) | (0.041) | (0.008) | (0.023) | (0.015) | (0.032) | (0.05) | (0.029) | (0.053) | (0.019) | (0.095) |
| Number of children | -0.209 | -0.078 | -0.063 | -0.07* | -0.023 | 0.161* | 0.237 | 0.053 | 0.071 | 0.089 | 0.057 | -0.282 |
| | (0.251) | (0.202) | (0.492) | (0.039) | (0.229) | (0.085) | (0.182) | (0.169) | (0.18) | (0.233) | (0.134) | (0.484) |
| Number of elderly | -0.39 | 0.336 | -0.096 | -0.194 | 0.044 | 0.062 | 0.297 | 0.061 | 0.582 | 0.849** | 0.128 | 1.059 |
| | (0.909) | (0.953) | (0.894) | (0.141) | (0.732) | (0.174) | (0.876) | (0.623) | (0.356) | (0.424) | (0.306) | (1.109) |
| Activity located at home | 1.642*** | 0.531 | 0.023 | 0.182** | 0.078 | 0.081 | 1.617 | -0.41 | 0.395 | 0.634 | 0.109 | -0.417 |
| | (0.317) | (0.349) | (0.898) | (0.079) | (0.524) | (0.141) | (0.996) | (0.642) | (0.418) | (0.399) | (0.238) | (0.856) |
| N | 407 | 560 | 588 | 377 | 685 | 278 | 1047 | 883 | 299 | 166 | 324 | 157 |

*Notes*: Table shows only the results of the inefficiency model, the estimates of the stochastic frontier production function are not presented. Controls not shown but included in the determinants of inefficiency equation: year indicators, human capital of the owner (age, experience, schooling, marital status), age of the IPU, factors 1,3,4 and 5 obtained using the PCA of household characteristics.

[1] It has been argued that interpersonal solidarity is in fact limited to the close family and neighbors and that the *fihavanana* viewed as a collective feeling of solidarity is a political discourse rather than a reality. Going further into this debate is beyond our scope, but it does signal the importance of solidarity in the Malagasy society, either as an observed or an imaginary practice (Wachsberger, 2009).

[2] In this study, family and kinship ties refer to any form of blood relationship. We also make use of the term social networks, which encompasses family and kinship ties, and more broadly community and ethnic ties. Family ties are the most proximate type of relationship while kinship ties are more distant and are characterized by socially recognized relationships based on supposed as well as actual genealogical ties (La Ferrara, 2007; Grimm et al., 2013).

[3] Both value added and capital are deflated using the Consumer Price Index (CPI).

[4] To save space, results of the PCA are available from the authors upon request.

[5] The survey used follows a commonly used categorization of the Malagasy population into 18 ethnic groups. The reality of this classification has been contested by historians who argue that it is a colonial invention, the result of "tribal" or "racial" views that prevailed among colonial ethnographers in the 1920s and 1930s. The resulting Merina versus côtiers (coastal people) opposition could have been used as a divide-and-rule colonial strategy. We use this classification because it is the most commonly adopted today and therefore the one for which data are available. The reader should keep in mind that the social organization of Madagascar is extremely complex, beyond the 18 ethnic groups. For example, there is a kind of "caste" system, inherited from the kingdoms existing prior to colonization, with many subdivisions, which can very roughly be grouped into three categories: the andriana (nobles), the hova (freemen) and the mainty (slaves) (Roubaud, 2000; Ramamonjisoa, 2002; Randrianja and Ellis, 2009; Jütersonke and Kartas, 2010).

[6] To calculate an (approximate) travel time, we multiplied the average automobile speed with the distance, taking into account the various types of roads (asphalt, dirt, river crossing, etc.) encountered to adapt the speed used.

[7] Regressions available from the authors upon request to save space.

[8] Unfortunately, the questionnaire used in 1995 did not include this second category, but only domestic tasks. Therefore, as the sample size is large enough, we will exclude the year 1995 from the analysis of the determinants of efficiency in Section 4.

[9] The dependency ratio is defined as the number of inactive over the number of active individuals in the household.

[10] See Aguirregabiria (2009) for a discussion of the various ways to correct the simultaneity bias.

[11] Another potential problem is that the multinomial logit may suffer from the Independence of Irrelevant Alternatives assumption (IIA), which in most cases is questionable. However, Bourguignon et al. (2007) argue that selection bias correction based on the multinomial logit model is a "reasonable alternative to multinomial normal models when the focus is on estimating an outcome over selected populations rather than on estimating the selection process itself. This seems even true when the IIA hypothesis is severely at odds". Since we are interested in results in the second stage regression, this allows us to be confident regarding the choice of a multinomial logit.

[12] In developing countries, females and children, for instance, may constitute a larger proportion of family than of hired labor. This would tend to drive down the marginal product of family relative to hired labor if the marginal productivity of women and children is lower than that of adult males (Deolalikar and Vijverberg, 1987). Besides, the skill differential between family and hired labor may also be an important source of heterogeneity in the productivity of workers.

[13] Moral hazard and their associated monitoring costs are mechanisms enlightening plausible greater productivity of family versus hired labor. Johnston and Le Roux (2007) report for instance that, for farmers, family labor can be more efficient than hired labor as it is assumed to be better incentivised and so will not shirk. The reason for superior incentives is that family labor will share the income generated by the farm as they may be "residual claimants to profits". Consequently, there will be shared incentives between entrepreneurs and workers (other household members) and so little need for additional supervision.

[14] Similarly, controlling for unobserved heterogeneity at the firm level with a panel for Vietnam, Nguyen and Nordman (2012) find evidence of greater elasticities of hired labor compared to family labor, in particular for informal businesses.

[15] The other regressors included generally have the expected sign: returns to the entrepreneur's formal education are positive and increasing with the level of education and the age of the firm has a positive effect once the entrepreneur's demographics are considered (Column 2). Lastly, the effect of the age of the owner is quadratic: increasing first until 31 years old, then decreasing. This is in accordance with the standard prediction of the human capital model stressing the existence of decreasing marginal returns to labor market experience.

[16] Non-zero returns to capital within the household can also occur in the presence of risk and risk aversion or of non-linearity in capital stocks (see Grimm et al., 2011).

[17] We thus confirm the finding of Nordman and Roubaud (2009) highlighting greater returns to education for women, using the same *1-2-3 Surveys* data for Madagascar in 1998, but extracting the information from the labor force module of the survey (Phase 1).

[18] To save on space, we only report selectivity-corrected estimates for the quantile regressions.

[19] The age of female-owned firms has a significant, positive effect above the 90$^{th}$ percentile, which may be the sign of endogenous exit, if larger firms survive longer than smaller firms. This would then tend to overestimate the effect of age on value added.

[20] The quantile regressions and the TRS are calculated on the full sample which includes an important number of IPUs that do not use all four types of inputs, in particular family and hired labor, as 70 percent of the sample are pure self-employment IPUs. Although we control for this by adding dummy variables for zero inputs, it is worth checking the robustness of our results to a restriction of the sample to IPUs with more than one worker. As the size of the sample is quite small, we ran quantile regressions at the 25$^{th}$, 50$^{th}$ and 75$^{th}$ percentiles and we merged family and hired labor into a single input. At the 25th and 75th percentiles, we still observe that the substitutability of the owner's labor is much lower in small female-owned enterprises than in larger ones, and the opposite pattern holds for male-owned IPUs (results not reported).

[21] We replicate the whole procedure 50 times to obtain standard errors. The precise methodology for constructing the counterfactual densities of the log performance at each considered quantile is described in Machado and Mata (2005).

[22] Although this paper focuses on the effect of these variables on technical efficiency, it may be of interest to look at another channel of transmission, through their impact on the level of inputs. We ran regressions for females and males separately with capital, owner labor and family labor as dependent variables. We did not find strong evidence of an effect of sharing norms and allocation of time within the household on the level of inputs of female and male entrepreneurs. One effect worth mentioning, however, is that the share of the same ethnic group in the neighborhood significantly increases the use of family labor in female-owned IPU, which would be consistent with the assumption that females are more subject to pressure to hire family workers by their local network. These regressions are available upon request from the authors.

[23] See Koopman (1951), Aigner et al. (1977), Meeusen and van den Broeck (1977), Battese and Coelli (1995).

[24] By definition, stochastic frontier analysis ignores distributional effects, in the sense that a firm is efficient as long as it is on its frontier of production, regardless of the position in the distribution. We are conscious that after having presented quantile regressions and decompositions this may pose a consistency problem. However, we attempt to take the problem into account by looking at various frontier models, including estimations by tercile of value added sub-samples.

[25] Other household composition variables were tested but were not significant. For the sake of simplicity we decided not to include them in the final model. We also investigated possible interacted effects of the household composition variables with domestic tasks but decided against keeping them in the model to preserve degrees of freedom, as they were not significant in most cases.

[26] Another possibility would be to estimate a unique frontier for both sexes and assess gender-differentiated inefficiency effects. However, the production functions estimated in the previous sections of the paper suggest that a common frontier for both male and female entrepreneurs is a very strong assumption, as we saw that returns to certain factors are gender-specific. Our approach is less normative in the sense that it does not assume that perfectly efficient male- and female-owned businesses would indeed perform equally (given their level of inputs). We are rather interested in how and why each sex is not able to reach his/her frontier, and evaluate whether certain constraints affect one more than the other, given a gender-specific frontier.

[27] Studies of the determinants of the inefficiency effects first used two-stage models, where the predicted efficiency term obtained with the stochastic frontier model was regressed of the explanatory variables of interest, as in Pitt and

Lee (1981). However two-stage models are biased because they first assume that the inefficiency effects are identically distributed in the stochastic frontier, then specify a regression model for the predicted efficiency firm, which contradicts the initial assumption (Battese and Coelli, 1995).

[28] Non-wage IPUs are businesses with more than one worker but no paid workers (mostly family members).

[29] Results not shown but available upon request from the authors.

[30] In the first two regressions (Table 10), we included the share of family labor in total labor, whose effect is not significant. As it is already captured in the frontier production function, we omitted it in the following regressions to preserve degrees of freedom.

[31] Whitehouse (2011) relates such a phenomenon in Bamako, as a taxi driver "would not accept a fare to his home neighborhood in that city, because he knew once he arrived there he was likely to be spotted by some relative who would insist on being driven somewhere for free."

[32] Source: 1-2-3 Surveys, Phase 1, 1998-2004, INSTAT-/DIAL/MADIO; author's calculation.

[33] From a point of view of family economics, where household members specialize either in market-oriented or in domestic activities to maximize the total welfare of the household, this would be a case of sub-optimal division of labor (Becker, 1981).

[34] The number of hours performing domestic tasks increases the efficiency of female entrepreneurs in the second tercile of value added. This effect is difficult to interpret, but it occurs in a category of businesses where children also increase efficiency, for which we propose several interpretations in the following paragraphs.