2020-08

Motherhood and Female Self-Employment: Theory and Evidence

Jean-Louis Bago
Sylvain Dessy

Mai/ May 2020

**Centre de recherche sur les risques
les enjeux économiques et les politiques publiques**

www.crrep.ca

ABSTRACT

Having young children can push women into self-employment to reconcile the demands of motherhood with their professional ambitions. However, if self-employment offers very few opportunities for decent pay and access to wage employment depends on education, then, motherhood may no longer matter to women's self-employment. In this paper, we formalize this idea theoretically and test it empirically using data from Nigeria. We use an identification strategy that corrects selection bias and the endogeneity of fertility jointly. We find no evidence of a causal effect of motherhood on women's self-employment. This result is robust to several alternative specifications. However, we also find that motherhood increases the probability of self-employment for single women but not for married women. These findings suggest that the social setting governs the importance of motherhood for women's self-employment.

JEL Classification: H52, I26, J21, J24, J46.

Keywords: Women, Motherhood, Marital status, Self-employment, Endogeneity, Selection bias.

Jean-Louis Bago : Department of Economics, Université Laval. CRREP.
Sylvain Dessy : Department of Economics, Université Laval. CRREP.

We are very grateful to Bruce Shearer, Bernard Fortin, Sonia Laszlo, and Luca Tiberti for helpful comments at various stages of this paper. We thank participants at the January 11th, 2019 CRREP/CDR conference held in Quebec City, Canada, and the 2019 meeting of the Canadian Economics Association held in Banff, Alberta, Canada.
Jean-Louis acknowledges financial support from *the Centre de recherche sur les risques, les enjeux économiques, et les politiques publiques* for the conduct of this project.

1 Introduction

In the labor literature, the primary explanation for women's participation in self-employment is the presence of young children (Connelly, 1992; Semykina, 2018). But the evidence supporting this explanation is documented only in developed countries, where women's probabilities of self-employment are much lower than those of men (Roche, 2014) and married couples post-marital living arrangements lead to the creation of nuclear families (Pew Research Center, 2019). In developed countries, post-marital living arrangements lead married couples to live on their own after marriage, forming nuclear families. This family structure places the burden of caring for young children solely on parents. By contrast, in many developing countries, post-marital living arrangements lead to the creation of multi-generation households[1] and extended family networks[2] (Pew Research Center, 2019). For example, in Nepal, 58% of households are extended family households. This figure also 58% in Liberia, 56% in Gabon, and 55% in Senegal. In Canada and the US, by contrast, the corresponding figures are significantly lower at respectively 9% and 11% (Pew Research Center, 2019). These facts on family structures around the world may represent a source of heterogeneity in the importance of motherhood to women's self-employment.

In this paper, we formalize this idea theoretically and then tests it using micro-level data from Nigeria. We begin by integrating women's marital status and post-marital living arrangements as variables in a Becker-type model of household occupational choice. In particular, we contrast nuclear and extended families. In our model, women either work for pay or do not work. Those who do not work for pay spend all their time caring for their young children. Those who do work, by contrast, can choose between wage employment and self-employment. Wage employment offers a better return on skills - as measured by education - than self-employment. This fact is the result of self-employment taking place predominantly in the informal sector where the use of rudimentary technologies is ubiquitous and positive network externalities in multi-employee workplaces (Kikuchi, 2007; Basker, 2012). A lower return on skills in self-employment thus represents a penalty that women face when they reject wage

[1] multi-generation households are households where children live with their parents and grandparents

[2] Extended family networks are understood as networks of adult relatives living near one another, with their families

employment. However, since children are not allowed in most wage employment settings, motherhood may increase the cost of wage employment for two essential reasons. First, the cost of outsourcing care for young children increases household expenses. Second, nonparental care may be an imperfect substitute for parental supervision, leading to lower well-being for children (Casarico and Sommacal, 2018).

From a theoretical point of view, therefore, both types of employment entail penalties for women. Which of the two kinds of work imposes the higher penalty on women may, therefore, depend on the social setting in which they live, as determined by their marital status and post-marital living arrangements. Indeed, we show that motherhood pushes single women into self-employment because all over the world, single working mothers tend to live on their own with their children (Pew Research Center, 2019). However, the situation is different for working married women. Where marriage leads to the emergence of either multi-generation households or extended family networks of relatives living close to one another, having young children does not influence women's self-employment. Instead, it is a lack of education that bars them from accessing better-pay wage employment. By contrast, where marriage leads to the creation of nuclear families, lack of access to affordable professional childcare services pushes women with young children into self-employment.

We take these predictions to the data and investigate the importance of motherhood for women's self-employment empirically. The challenge in identifying the causal effect of motherhood on women's self-employment probabilities stems from the potential presence of two problems—first, the nonrandom self-selection of working-age women in the paid labor force and, second, the endogeneity of fertility decisions. As in Semykina (2018), we address these two problems jointly. However, to correct for the endogeneity of the motherhood variable, we follow Aguero and Marks (2008) and Agüero and Marks (2011) in instrumenting fertility by infertility shocks. Consistent estimates result from the application of a partial maximum likelihood estimator's (MLE) method, as in Semykina (2018).

Our data comes from the 2018 *Demographics and health surveys* for Nigeria. The choice of Nigeria stems from the fact that, like most sub-Saharan African countries, it features different family structures, including nuclear families, as well as networks of extended family members living either in multi-generation households or near one another. For example, in Nigeria, the

proportion of multi-generation households— with both children under the age of 15 and family members aged 60 or higher— is 14%. In comparison, the corresponding figure is a mere 2% in the United States (United Nations, 2018).

When we correct for nonrandom self-selection of women into the paid labor force and endogeneity of fertility decisions jointly, we find no evidence of a causal effect of motherhood on women's self-employment probabilities in Nigeria. By contrast, education reduces these probabilities. Having at least a secondary education reduces a woman's likelihood of self-employment by roughly 9 percentage points. Furthermore, as a surprising result, we find that motherhood pushes women into the paid labor force. Indeed, having at least one preschool child increases the probability that a woman works for pay by 15 percentage points. The existing literature finds mixed results about the motherhood effect of women's participation in the paid labor force (Aguero and Marks, 2008; Agüero and Marks, 2011; Agüero et al., 2012; Semykina, 2018). Our theory explains this surprising result by the struggle to survive, which characterizes household economic conditions in many sub-Saharan African countries, including Nigeria. This struggle to survive forces both parents to work for pay, to meet their basic subsistence needs.

We show that all the findings above are robust to the use of an alternative estimation method. Both the partial MLE (Semykina, 2018) and the two-stage residual inclusion method (Schwiebert, 2015) give similar results, with almost identical magnitudes.

However, when we stratify our sample by marital status to account for the fact that single mothers tend to live on their own with their children, we uncover a different reality. Motherhood matters to the self-employment of single women but not married women. Indeed, having at least one preschool child increases the probability that a single working woman is self-employed by 17 percentage points. This effect is significant at the 1% level. On the other hand, we find no evidence that motherhood pushes married women into self-employment. This finding stems from the fact that, as post-marital living arrangements, both nuclear and extended family structures coexist in Nigeria.

Overall, the main takeaway from this study is that the reported positive effect of motherhood on women's self-employment probabilities is not universal. The importance of young children for women's self-employment depends not only on the context in which this work

takes place but also on women's marital status and post-marital living arrangements. In many developing countries, and unlike in developed countries, self-employment is predominantly an activity of the informal sector characterized by reliance on rudimentary technologies. By contrast, wage employment takes place mainly in the formal sector, where education is an essential requirement for finding a job. In this socio-economic context, women's self-employment probabilities may be driven primarily by low educational attainments and households' struggle to survive. Our findings thus indicate that public policy aimed at pulling developing countries' women out of low-pay self-employment should prioritize childcare supports for single mothers as well as investment in female education.

Our study contributes to the literature linking women's characteristics to their labor market outcomes (Aguero and Marks, 2008; Agüero and Marks, 2011; Agüero et al., 2012; Heath, 2017; Semykina, 2018). Aguero and Marks (2008) use infertility shocks as a source of exogenous variation in fertility outcomes to analyze the causal effect of motherhood on women's labor market participation in the Latin American context. They find no evidence that children have a causal impact on women's labor force participation. However, their definition of labor market participation includes involvement in both paid and unpaid work. Agüero and Marks (2011) also use infertility shocks to study the causal effect of childbearing on women's employment outcomes in the context of 26 developing countries. They find that the presence of children limits women's opportunities for paid employment. Our paper differs from those of Aguero and Marks (2008) and Agüero and Marks (2011) by going beyond women's labor force participation decisions to include the choice between self-employment and wage employment among working women.

Other papers that examine the causal effect of motherhood on women's self-employment probabilities are Connelly (1992), Bianchi (2000), Wellington (2006), Boserup et al. (2013), Semykina (2018), Lim (2018). But our study more closely related to Semykina (2018). In her paper, Semykina (2018) studies the effect of young children on women's self-employment probabilities in using US data. After accounting for selection bias and endogeneity jointly, she finds that motherhood increases the self-employment probabilities of women, while education does not matter for women's self-employment. Our study focuses on Nigeria— a developing country where evidence shows that employment growth among women has been driven by the labor

force participation of women with the lowest levels of educational attainments (World Bank, 2015).

The rest of this paper is structured as follows. Section 2 presents the theoretical model from which we derive some testable predictions. Section 3 examines the data and presents the stylized facts while Section 4 presents the empirical model. Section 5 presents the empirical results. Robustness check analyses are presented in section 6 and Section 7 concludes.

2 Theoretical Model

In this section, we develop a simple model to capture the relationship between a woman's individual characteristics, the social setting in which she lives and her employment decisions to motivate the empirical analysis to follow.

The main features of our model draw from a number of existing household models of occupational choice, including Fernández et al. (2004), De la Croix and Doepke (2004), Tertilt (2006) and Bertrand et al. (2016). We adapt these existing models to fit a variety of social settings, and generate predictions that are context-specific.

2.1 Fundamentals

There is a measure one of households indexed by $i \in [0,1]$. A typical household in this environment is identified by a vector containing the woman's individual characteristics $X_i = \left(E_i, K_5^i, M_i\right)$, where $E_i \geqslant 0$ denotes her level of education, $M_i \in \{0,1\}$, her marital status, and $K_5^i \geqslant 0$, the number of preschoolers she has.

We classify households into two types: single-parent households, headed by a woman ($M_i = 0$), and two-parent households, co-headed by a man and a woman ($M_i = 1$). These are the dominant family types observed in the world (*World Family Map 2015*)[3]. Furthermore, we take the intra-household division of labor as exogenously given. According to this division of labor, working mothers with preschoolers face a time-constraint for their participation in the labor force. This induces co-headed households (those in which the woman has marital status

[3]See the WORLD FAMILY MAP 2015. *Mapping Family Change and Child Well-Being Outcomes*. Available online at https://ifstudies.org/ifs-admin/resources/reports/wfm-2015-forweb.pdf

$M_i = 1$) to choose between a traditional single-earner family model ($d_i = 0$) and a dual-earner family model ($d_i = 1$) to support the family. In other words, a married woman's labor force participation decision is intertwined with the household choice of family model. There is no involuntary unemployment among women in this environment.

In a co-headed household, the male spouse's income denoted as θ_i is exogenously given, and thus is a contributing source of heterogeneity across women. A co-headed household in which the woman does not participate in a paid employment is one that adopted the family strategy $d_i = 0$, whereby the wife fully specializes in care-giving and homemaking at the expense of paid employment. However, for a co-headed household that chooses the dual-earner family model whereby both spouses engage in paid employment ($d_i = 1$), there is a choice to be made between self-employment ($y_i = 0$) and wage employment ($y_i = 1$). Self-employment allows a woman with young children to care for her children while working. Wage employment, by contrast, enables her to fully specialize in paid employment at the expense of family responsibilities, provided she can find suitable childcare arrangements for her preschoolers, either free of charge if living in an extended family network or for a fee if living in a nuclear family.

We assume that a woman's level of education, E_i, and the number of preschoolers she has, K_5^i, are exogenously given. This assumption allows us to focus on deriving the causal effects these two characteristics may have on women's employment outcomes conditional on the marriage custom, if married. We denote as $S_i \in \{0, 1\}$, the marriage custom practiced by the community in which a woman lives. $S_i = 0$ corresponds to settings where marriage leads to the creation of a network of extended family members, while $S_i = 1$ corresponds to settings where marriage leads to the creation of a nuclear family.

Note that we restrict attention to the presence of preschoolers in the household, ignoring school-age children. This is to emphasize the need to care for preschoolers. Evidence shows that even in developing countries enrolment in primary education is near universal (UNDP's *Human Development Report 2016*), implying that mothers of school-age children are relatively free to participate in the paid labor force.

2.2 Occupation-Specific Time Allocation Constraint

The decision on the woman's occupational profile, $(d_i, y_i) \in \{0,1\} \times \{0,1\}$, is made by the household as a unit, to maximize an appropriately defined household utility function. A single woman has no choice but to participate in the paid labor force, to support herself and her children, if any. Her only decision, therefore, concerns the type of employment to sort in to $y_i \in \{0,1\}$. We will return to this decision further below.

Turning to a co-headed household, suppose it chooses the occupational profile $(d_i, y_i) = (0, y_i)$ for the woman. Then, the husband works for pay and the wife fully specializes in caregiving and homemaking, and so does not participate in the paid labor force. Such a woman is totally dependent on her husband for her needs and those of the children they have. Next, suppose that this co-headed household i chooses the occupational profile $(d_i, y_i) = (1,1)$ instead. Then, both spouses participate in the paid labor force ($d_i = 1$), but the wife is self-employed ($y_i = 1$). If this household chooses the occupational profile $(d_i, y_i) = (1,1)$ for the woman, then, again, both spouses participate in the paid labor force ($d_i = 1$), but the wife is wage-employed ($y_i = 1$).

Wage-employed mothers cannot work while caring for their children. Wage-employed mothers single mothers ($M_i = 0$)) or wage-employed married mothers ($M_i = 1$)) under the neolocal marriage custom $S_i = 1$) must pay for childcare services from professionals caregivers during the time they are at their paid job. By contrast, a married mother ($M_i = 1$)) living in a setting $S_i = 0$) pays no child care fees, as her children are being watched by members of the extended family network.

2.3 Preferences and Budget Constraints

Each household derives utility from the spouses' joint consumption of a numeraire, c_a^i and from a household public good denoted as G_i.[4] To motivate the desire for children, assume that not having offspring generates a disutility $\delta > 0$. Then, a household utility payoff writes as follows:

$$u_i = \log c_i^a + \lambda\gamma \log G_i, \tag{1}$$

[4] In a household headed by a single woman, c_i^a denotes her own consumption.

where $\gamma > 0$, and λ denotes an indicator function equal to 1 if the household has at least one child (i.e., $K_5^i > 0$) and 0 otherwise.

The level of public good enjoyed by household i, G_i, is a composite of the quantity and quality of children (Becker and Lewis, 1973; Becker and Tomes, 1976). We follow the literature in assuming that the production technology for the household public good is Cobb-Douglas in the quantity and quality of children (De La Croix and Doepke, 2003; De la Croix and Doepke, 2004).[5] A household i with K_5^i young children nurtured into achieving each a level of quality, q_i, generates a level of household public good given by:

$$G_i = (q_i)^{\beta}(K_5^i)^{1-\beta}, \tag{2}$$

where $\beta \in (0,1)$. In Equation (2), by quality of offspring we mean the level of cognitive and socioemotional development attained by a child. Consistent with the literature on child development (Becker and Tomes, 1976; Currie and Moretti, 2003; Del Boca et al., 2013; Heckman and Mosso, 2014; Heath, 2017), we model child quality as produced through a Cobb-Douglas function of monetary and effective time inputs. On the one hand, household income is necessary to meet the nutritional needs of children. On the other hand, caregiving to young children plays an important role in their cognitive and socioemotional development, with long-term implications for their human capital outcomes. We assume that the effect of caregiving on child quality, q_i, depends not only on the effective time spent interacting with children, which is the hours spent interacting with a child times the mother's level of education, E_i. Therefore, in a single-earner household—i.e., one in which the mother specializes completely in caregiving to her young children—, child quality is given by:

$$q_i = (1 + E_i)^{1-\eta} \left(c_i^k\right)^{\eta}, \tag{3}$$

where $\eta \in (0,1)$. Child quality increases with mother's education because maternal education impacts the quality of the interactions a mother has with her young children, which, in turn, affects their cognitive and socioemotional development.

A self-employed mother (i.e., $y_i = 1$) interacts with her young children while working.

[5] We ignore school-age or grown up children for simplicity.

Arguably, this lack of specialization may adversely impact the transmission of human capital from mother to children, due to either fatigue or preoccupations with the outcome of business operations. For a self-employed mother of K_5^i preschool children, therefore, child quality is given by:

$$q_i = (1 + \delta_1 \, E_i)^{1-\eta} \left(c_i^k\right)^{\eta}, \tag{4}$$

where $\delta_1 \in [0, 1]$ measures the marginal productivity of the mother's knowledge in her limited interactions with her young children while self-employed. As long as $\delta_1 < 1$, lack of maternal specialization in caregiving to her young children adversely impacts their cognitive and socioemotional development. By contrast, $\delta_1 = 1$ means there is no such adverse effect.

A wage-employed mother (i.e., $y_i = 0$), does not interact with her young children for many hours during a day of work. But she uses her knowledge (as measured by her level of education, E_i) to secure a stimulating childcare arrangement for her young children, either free of charge through the extended family network or through professional childcare providers for a fee in a nuclear family setting. In the literature on child development, there are concerns that a child's separation from the primary caregiver— the mother—, may trigger stress and anxiety in the child, with adverse effects on his or her socioemotional development (Mercer, 2006). Echoing these concerns, we model the impact of a wage-employed mother's limited interactions with her young children as measured by $\delta_0 \in [0, 1]$. Without loss of generality, we assume that caregiving in the extended family and in a professional childcare are perfectly substitutable. Consequently, child quality for a wage-employed mother is given by:

$$q_i = (1 + \delta_0 \, E_i)^{1-\eta} \left(c_i^k\right)^{\eta}. \tag{5}$$

When $\delta_0 = 1$, parental and nonparental caregiving are perfect substitutes in the production of child quality.

Assumption 1. $\delta_0 < \delta_1 \leqslant 1$,

Assumption 1 ensures that there is a balancing act motive for women's self-employment. It is important to note, however, that the literature addressing parental versus nonparental caregiving in relation to child development yields no consensus on the ranking of these caregiving

strategies. Some studies find that nonparental childcare arrangements adversely impact child development when the quality of such care is poor (Baker et al., 2008; Herbst, 2013; Fort et al., 2017), while others find that it boosts child development where the quality is high (Gupta and Simonsen, 2010; Noboa-Hidalgo and Urzua, 2012; Drange and Havnes, 2014). Moreover, with respect to the role of parental interactions for children's long-term development, several recent studies find no effect (Rasmussen, 2010; Dustmann and Schönberg, 2012; Baker and Milligan, 2015), while several others find some positive effects on child human capital (Carneiro et al., 2015; Danzer and Lavy, 2018; Drange and Havnes, 2019). We discuss our theoretical results in light of this lack of consensus in the literature on parental vs. nonparental caregiving to preschool children.

As in Fernández et al. (2004) and Bertrand et al. (2016), in a dual-earner household, spouses pool their individual incomes. We denote as $R\left(d_i, y_i, X_i\right)$ total household i income, where $X_i = \left(E_i, K_5^i, M_i, S_i\right)$ denotes the vector of the woman's individual characteristics, for all $i \in [0,1]$. We denote a woman's occupation-specific income as $W\left(y_i, X_i\right)$. A household i's occupation-specific income thus writes as follows:

$$R\left(d_i, y_i, X_i\right) = \theta_i M_i + d_i\left[W\left(y_i, X_i\right) - \left(1 - y_i\right)\left(\underline{c} - \epsilon E_i\right)\right], \tag{6}$$

for all $\left(d_i, y_i, X_i\right)$, where θ_i denotes the male partner's exogenously given income, $\underline{c} - \epsilon E_i$, the search cost a woman with level of education E_i must incur to find a wage employment, $\underline{c} > 0$ is scalar and $\epsilon \in (0,1)$. In other words, the cost of searching for wage employment a woman is higher the less educated the individual. This implies that low education is a barrier to wage employment. Support for this assumption derives from two pieces of evidence. First, self-employment developing countries takes place predominantly in the informal sector while wage employment takes place predominantly in the formal sector (Conference, 2013). Second, informal work these countries is often undertaken because of barriers to entry to formal work where education requirements are high (Conference, 2013).

A woman with individual characteristics X_i who works for pay ($d_i = 1$) earn an income whose level depends on the type of employment she holds, y_i. The income of a self-employed woman ($y_i = 1$) is generated from the operation of a constant-return-to scale technology that

requires effective labor only, which is the product of a woman's raw labor time, in total quantity normalized to unity, and her level of education, E_i. We assume that an hour of work in self-employment yields an income $\underline{A}E_i$ for a self-employed woman with no young children, where $\underline{A} > 0$ is a productivity factor. This modelling assumption is consistent evidence showing that the return to self-employment increases with the worker's level of education (Moock et al., 1989). The labor income of a self-employed woman with a vector of individual characteristics X_i thus is $W(1, X_i) = \underline{A}E_i$.

In wage employment, a woman also allocates her entire unit endowment of time to paid work and provides no care-labor. We assume that the production technology in wage employment is also constant-return-to scale to effective labor. An hour of work in wage employment yields an income equal to AE_i. This modelling assumption is consistent with empirical evidence showing that the return to wage employment rises with a worker's level of education (Garcia-Mainar and Montuenga-Gomez, 2005).

Assumption 2. $A > \underline{A}$

Assumption 2 states that effective labor is more productive in wage employment than in self-employment. In other words, there is a productivity premium for wage employment, and this premium rises with a woman's level of education, E_i. This assumption can be justified by the fact that a modern workplace functions like a network of intertwined operations, with an optimal division of complementary tasks among its members (employees). This creates network externalities (Kikuchi, 2007; Basker, 2012) that positively impact individual productivity among workers.

A wage-employed woman has a level of income given by $W(0, X_i) = AE_i$. More generally, we can rewrites the employment-specific income of a female worker as follows:

$$W(y_i, X_i) = \begin{cases} AE_i & \text{if } y_i = 0 \\ \\ \underline{A}E_i & \text{if } y_i = 1 \end{cases} \tag{7}$$

for all i such that $d_i = 1$.

The budget constraint faced by a household with a voluntarily unemployed woman ($d_i = 0$) is given by $c_i^a + c_i^k K_5^i \leqslant R(0, y_i, X_i) \equiv \theta_i$, and that of a household with a self-employed woman

is $c_i^a + c_i^k K_5^i \leqslant R\left(1, 1, X_i\right)$.

Consider a household with K_5^i preschool children living in a community that practices the marriage custom S_i. suppose the the woman is wage-employed, and denote as p_c the per child cost of an hour of childcare, then total childcare expenditures amount to $p_c S_i K_5^i$. In other words, for a co-headed household living in a setting where extended family networks are common (i.e., $S_i = 0$), total childcare expenditures amount to zero.

The budget constraint of a household with a wage-employed woman thus is $c_i^a + c_i^k K_5^i + p_c S_i K_5^i \leqslant R\left(1, 1, X_i\right)$, due to childcare costs. More generally, the budget constraint of household i with a woman X_i whose occupational profile is $(d_i, y_i) \in \{0, 1\} \times \{0, 1\}$ obtains as follows:

$$c_i^a + c_i^k K_5^i + \left(1 - y_i\right) d_i p_c K_5^i S_i \leqslant R\left(d_i, y_i, X_i\right), \tag{8}$$

for all i. In other words, a self-employed mother (i.e., $y_i = 1$) pays no child care fee, as does a wage-employed mother (i.e., $y_i = 0$) in the setting $S_i = 0$.

Given X_i, each household i has up to three sequential decisions to make. First, the household decides on the woman's labor force participation $d_i \in \{0, 1\}$. This decision is akin to choosing a family model (either a single-earner, or a dual-earner, model). A co-headed household in which the woman does not participate in the paid labor force ($d_i = 0$) is one that opted for the single-earner model. For this type of household, the next decision is on how to allocate its total income across household expenditures. By contrast, a dual-earner household ($d_i = 1$) must next chooses the type of employment for the female spouse: self-employment ($y_i = 1$) or wage employment ($y_i = 0$). Based on the outcome of this employment type decision, household i then decides on the allocation of its total income across household expenditures.

2.4 Occupation-Specific Payoffs

We next characterize household i's utility payoff associated with each occupational profile $(d_i, y_i) \in \{0, 1\} \times \{0, 1\}$, for the woman. We denote as $U^i\left(d_i, y_i, X_i\right)$ the utility payoff of a household in which the woman has an occupational profile (d_i, y_i), when her individual characteristics are given by X_i. For the time being, assume $\lambda = 1$, i.e., all households have young children.

2.4.1 The Non Participation Payoff

An unemployed woman has no income of her own. Therefore, the decision not to participate in the paid labor force is only relevant for married women. Consequently, we assume that all women not involved in paid work are married ($M_i = 1$), and can be supported by their spouses.

From (1), substituting in (2), (3), (6), and (8), yields the utility payoff of a single-earner, co-headed, household ($d_i = 0$), as follows:

$$U^i(0, y_i, X_i) = \log\left(\theta_i - c_i^k K_5^i\right) + \eta\gamma\beta \log\left(c_i^k\right) + \underline{\beta} \log(1 + E_i) + \gamma(1 - \beta) \log K_5^i, \tag{9}$$

for all i, such that $(d_i, y_i) = (0, y_i)$, where $\underline{\beta} = (1 - \eta)\gamma\beta$. For this household, the problem is to choose c_i^k to solve the following problem:

$$\max_{c_i^k} U^i(0, y_i, X_i).$$

The interior solution to this problem is

$$c_i^k = \frac{\eta\gamma\beta\theta_i}{(1 + \eta\gamma\beta) K_5^i} \tag{10}$$

for all i such that $(d_i, y_i) = (0, y_i)$. Expression (10) illustrates the well-known quantity-quality trade-off (Becker and Lewis, 1973), whereby having more children (i.e., a higher K_5^i) reduces the quality each child can attain.

Substituting (10) back into (9), rearranging terms, yields the optimal utility payoff of non participation in the paid labor force as follows:

$$U^{i*}(0, y_i, X_i) = \bar{\beta} \log(\theta_i) + \underline{\beta} \log(1 + E_i) + \gamma[1 - \beta(1 + \gamma\eta)] \log K_5^i + Z, \tag{11}$$

where $\bar{\beta} := 1 + \gamma\beta\eta$ and

$$Z_0 := \gamma\eta\beta \log \gamma\eta\beta - (1 + \gamma\eta\beta) \log(1 + \gamma\eta\beta).$$

Observe from expression (11) that unless households in this environment have a preference bias

towards quantity, relative to quality, of offspring, i.e.,

$$\beta < 1/(1+\gamma\eta), \tag{12}$$

having children would reduce a household well-being. Furthermore, the payoff of non participation rises with the female spouse's level of education: $\partial U^{i*}(0, y_i, X_i)/\partial E_i > 0$. This is because a mother's level of education is a contributing factor to the quality of children she has. Furthermore, if condition (12) holds, then motherhood (i.e., having children) increases the utility payoff of non participation in the paid labor force: $\partial U^{i*}(0, y_i, X_i)/\partial K_5^i > 0$. This effect stems simply from the fact that as a unit, spouses have a preference bias towards quantity relative to quality of offspring.

Importantly, expression (11) implies that a married woman who receives a transfer θ_i such that $\theta_i \leqslant \underline{c} K_5^i$ will have no choice but to work for pay, otherwise her children will not survive. This survival motive for women's employment, where ever prevalent, would contradict the presumption that marriage is a barrier to women's employment.

2.4.2 The self-employment Payoff

Next, we characterize the utility payoff of a household in which the woman is self-employed. Suppose household i chooses the play of the sequential strategy profile $(d_i, y_i) = (1, 1)$. According to this profile, a woman (married or not) participates in the paid labor force ($d_i = 1$), earning an income as self-employed worker ($y_i = 1$). Using (6) and (7), yields this household's total income as follows:

$$R(1, 1, X_i) = \theta_i M_i + \underline{A} E_i. \tag{13}$$

From (1), substituting in (2), (3), (8), and (13), yields the utility payoff of this household as follows:

$$\begin{aligned} U^i(1, 1, X_i) = \log\left[\theta_i M_i + \underline{A} E_i - c_i^k K_5^i\right] + \gamma\beta\eta \log\left(c_i^k\right) \\ + \underline{\beta} \log(1 + \delta_1 E_i) + \gamma(1-\beta) \log K_5^i, \end{aligned} \tag{14}$$

for all i such that $(d_i, y_i) = (1, 1)$. A typical household i for which the woman is self-employed

thus solves the following problem:

$$\max_{\langle c_i^k \rangle} U^i\left(1,1,X_i\right),$$

for all i such that $(d_i, y_i) = (1,1)$. The interior solution to this problem writes as follows:

$$c_i^k = \frac{\gamma\beta\eta}{1+\gamma\beta\eta}\left[\frac{\theta_i M_i + \underline{A}E_i}{K_5^i}\right] \tag{15}$$

From (14), substituting in (15), rearranging terms, yields the optimal payoff of self-employment as follows:

$$\begin{aligned} U^{i*}\left(1,1,X_i\right) = {} & \bar{\beta}\log\left(\theta_i M_i + \underline{A}E_i\right) \\ & + \underline{\beta}\log\left(1+\delta_1\ E_i\right) + \gamma\left[1-\beta\left(1+\gamma\eta\right)\right]\log K_5^i + Z, \end{aligned} \tag{16}$$

Observe from (16) that

$$U^{i*}\left(1,1,X_i\right)|_{M_i=1} - U^{i*}\left(1,1,X_i\right)|_{M_i=0} > 0,$$

implying that the payoff of self-employment is higher for a married, than for a single, woman. Furthermore, partial differentiation of (16) shows that $\partial U^{i*}\left(1,1,X_i\right)/\partial E_i > 0$. Being more educated increases the payoff of self-employment. This is because education increases the productivity of a self-employed woman. Likewise, $\partial U^{i*}\left(1,1,X_i\right)/\partial K_5^i > 0$ owing to parents' preferential bias towards quantity of offspring.

2.4.3 The Wage Employment Payoff

To be wage-employed, a woman must first incur a search cost, $\underline{c} - \epsilon E_i$. A household in which the female spouse is wage-employed has a profile $(d_i, y_i) = (1,0)$. According to this profile, the woman participates in the paid labor force ($d_i = 1$), earning an income as wage-employed worker ($y_i = 0$). Combining (6) and (7) yields the total income of such a household as follows:

$$R\left(1,0,X_i\right) = \theta_i M_i + AE_i - \left(\underline{c} - \epsilon E_i\right) \tag{17}$$

for all i, such that $(d_i, y_i) = (1, 0)$.

From (1), substituting in (2), (3), (8), and (17) yields the utility payoff of wage employment as follows:

$$U^i(1, 0, X_i) = \log\left(\theta_i M_i + AE_i - (\underline{c} - \epsilon E_i) - p_c K_5^i S_i - c_i^k K_5^i\right) + \gamma\beta\eta \log\left(c_i^k\right) + (1 - \eta)\gamma\beta \log(1 + \delta_0 E_i) + \gamma(1 - \beta) \log K_5^i \quad (18)$$

for all i, such that $(d_i, y_i) = (1, 0)$.

This household's problem thus reduces to:

$$\max_{\langle c_i^k \rangle} U^i(1, 0, X_i).$$

The interior solution to this problem writes as follows:

$$c_i^k = \frac{\gamma\eta\beta}{(1 + \beta\gamma\eta)K_5^i}\left[\theta_i M_i + AE_i - (\underline{c} - \epsilon E_i) - p_c K_5^i S_i\right] \quad (19)$$

From (18), substituting in (19), rearranging terms yields the utility payoff of wage employment as follows:

$$U^{i*}(1, 0, X_i) = \bar{\beta} \log\left[\theta_i M_i + AE_i - (\underline{c} - \epsilon E_i) - p_c K_5^i S_i\right] + \underline{\beta} \log(1 + \delta_0 E_i) + \gamma\left[1 - \beta(1 + \gamma\eta)\right] \log K_5^i + Z \quad (20)$$

for all i such that $(d_i, y_i) = (1, 0)$.

Just as in the case of self-employment, marriage has a positive effect on the payoff of wage employment,

$$U^{i*}(1, 0, X_i)\,|_{M_i=1} - U^{i*}(1, 0, X_i)\,|_{M_i=0} > 0,$$

as does education, $\partial U^{i*}(1, 1, X_i)/\partial E_i > 0$. Furthermore, a close inspection of expression (20) reveals that motherhood has an ambiguous effect on this payoff, due to both childcare expenditures and nutrition costs.

2.5 Optimal Occupational Decisions

As stated above, each household must choose the occupation profile of the woman, $(d_i, y_i) \in \{0,1\} \times \{0,1\}$, to maximize its utility payoff:

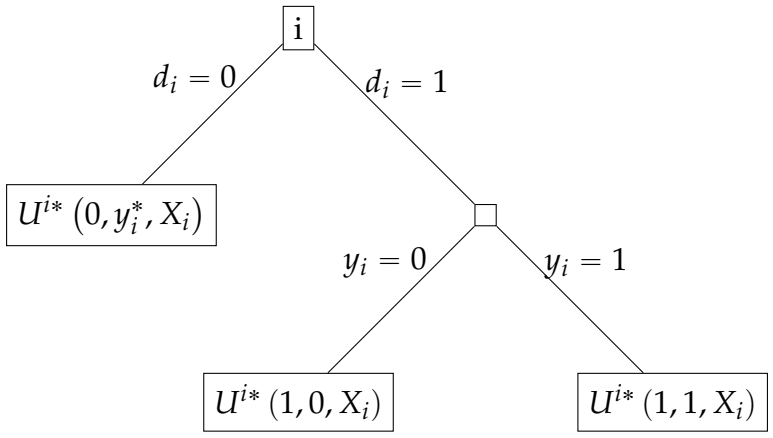
$$\max_{d_i} \left\{ \max_{(y_i)} U^{i*}(d_i, y_i, M_i) \right\}$$

where

$$U^{i*}(1, y_i, X_i) = \begin{cases} U^{i*}(1,0,X_i) & \text{if } y_i = 0 \\ \\ U^{i*}(1,1,X_i) & \text{if } y_i = 1 \end{cases}.$$

The decision tree characterizing this two-stage problem is represented below (Figure 1), along with the associated utility payoffs:

Figure 1: *Decision tree*



We apply a backward induction process to solve this two-stage problem. First, given that in the first stage the woman plans to be employed, in the second stage the household then must choose the type of employment this woman will perform,y_i: this amounts to solving the following problem:

$$\max_{(y_i)} U^{i*}(1, y_i, M_i).$$

Next, given the optimal decision at this second stage, y_i^*, the household then decides whether or not the woman will participate in the paid labor force. This amounts to solving the following

decision problem:

$$\max_{d_i} \hat{U}^{i*}\left(d_i, y_i^*, M_i\right),$$

where $y_i^* \equiv \arg\max_{y_i} U^{i*}(1, y_i, M_i)$. We therefore solve the second stage first. We formally apply this backward induction process in what follows.

2.5.1 Optimal Choice of Employment Type

Given that the woman in household i plans to be employed, what type of paid employment will her household select? Since selection of employment type is a binary decision in our model, household i will make this selection by balancing between the utility payoff of self-employment and that of wage employment.

Denote as $Y_i^* := \hat{U}^i(1, 1, X_i) - \hat{U}^i(1, 0, X_i)$, the net payoff gain of selecting self-employment as the female spouse's paid employment. Since theory does not put any restriction on δ_1 and δ_1, without loss of generality, let us assume that $\delta_1 - \delta_1 \to 0$. Then, using (16) and (18), this net payoff rewrites as follows:

$$Y_i^* = \bar{\beta} \log\left[\frac{\theta_i M_i + \underline{A} E_i}{\theta_i M_i + \bar{A} E_i - \bar{c} - p_c K_5^i S_i}\right] = H(E_i, K_5^i, M_i, S_i), \tag{21}$$

where $\bar{A} = A + \epsilon$. Self-employment (i.e., $y_i = 1$) is optimal when the female spouse has individual characteristics X_i such that

$$Y_i^* > 0. \tag{22}$$

By contrast, wage employment (i.e., $y_i = 0$) is optimal when these individual characteristics are such that

$$Y_i^* < 0. \tag{23}$$

(i) Does motherhood drive women's self-employment?

To address this issue, assume that K_5^i is a continuous variable, so that Y_i^* is differentiable in K_5^i. The structure (21) of enables us to distinguish two cases.

First, when $S_i = 1$, i.e., a nuclear family structure is prevalent, then a simple inspection of expression (21) shows that Y_K^* is strictly increasing in the number of preschool children a woman has. This implies that motherhood promotes married women's self-employment in social settings characterized by the prevalence of nuclear families. This case is consistent with empirical evidence from developed countries showing that motherhood has a positive causal effect on women's self-employment probabilities (Semykina, 2018).

By contrast, when $S_i = 0$, corresponding to a social setting characterized by the prevalence of extended family networks then, it follows from expression (21) that Y_K^*, in this context, is independent of the number of young children a woman has. In other words, in a social setting characterized by the prevalence of extended family networks (i.e., $S_i = 0$), motherhood does not matter to women's self-employment decisions.

The following proposition therefore summarizes the above discussion of the effect of motherhood on women's self-employment decision:

Proposition 2.1 *Let Assumptions 1 - 2 hold simultaneously. Then,*
(i) in social settings characterized by the prevalence of extended family networks (i.e., $S_i = 0$), motherhood does not matter to women's self-employment decision.
(ii) By contrast, in communities social settings characterized by the prevalence of nuclear families (i.e., $S_i = 1$), motherhood pushes women into self-employment.

Proposition 2.1 states that for women, the social setting determining their post-marital living arrangements governs the importance of motherhood to self-employment decisions. These results suggest that public policy toward women's self-employment should factor in the social setting of this type of occupation.

(ii) Does marriage influence women's self-employment?

To explore the effect of marriage on women's self-employment, we compare the net payoff of self-employment for a married woman ($M_i = 1$) to the corresponding net payoff for a single, but otherwise identical, woman ($M_i = 0$). For this purpose, we compute the difference $\Delta Y_M^* = H(E_i, K_5^i, 1, S_i) - H(E_i, K_5^i, 0, S_i)$, using (21). We can then show that this difference is equivalent

to:

$$\Delta Y_M^* = \bar{\beta} \log \left[\frac{\theta_i + \underline{A} E_i}{\underline{A} E_i} \right] - \bar{\beta} \log \left[\frac{\theta_i + \bar{A} E_i - \bar{c} - p_c S_i K_5^i}{\bar{A} E_i - \bar{c} - p_c K_5^i} \right]. \tag{24}$$

Under Assumptions 1 and 2, the sign of (24) is clearly ambiguous, as it depends on the woman's other characteristics (including her level of education and the number of children she has), as well as the characteristics of the socioeconomic environment in which she lives. In other words, the question of whether marriage promotes women's self-employment is purely an empirical issue.

(iii) Does education influence women's self-employment?

To clarify as much as possible the picture of the effect of education on women's self- employment, we assume that education is a continuous variable and take the partial derivative of expression (21) with respect to E_i. For a married woman (i.e., $M_i = 1$), this yields:

$$H_E = -\bar{\beta} \left(\frac{\theta_i \left(\bar{A} - \underline{A} \right) + \underline{A} \left(\bar{c} + p_c K_5^i S_i \right)}{\left[\theta_i + \underline{A} E_i \right] \left[\theta_i + \bar{A} E_i - \bar{c} - p_c K_5^i S_i \right]} \right) \tag{25}$$

Under Assumptions 1-2, the partial derivative in (25) is strictly negative. Hence the following proposition:

Proposition 2.2 *Let Assumptions 1-2 hold simultaneously. Then, education has a negative effect on women's participation in self-employment.*

Proposition 2.3 stems from the fact that there is a productivity premium for education in wage employment. In other words, the return to education is higher in wage employment than in self-employment. This fact provides educated individuals with the incentive to pursue wage employment, to cash in on this productivity premium.

Note that all the above propositions are conditional on women's decision to participate in the paid labor force. In particular, Proposition 2.2 is a prediction on how working women sort into different types of employment on the basis of their respective levels of education, but not necessarily a prediction on how women in general behave with respect to this issue, including

those who chose not to participate in the paid labor force. Hence, our focus on the labor force participation decision.

2.5.2 Optimal Participation Decision

What factors promote women's participation in the labor force? Are these factors identical to those that influence the type of employment they engaged in? Does the social setting matters to the answers to these questions? We address these issues by analyzing the first stage labor force participation decision problem facing each household. This decision problem rewrites as follows:

$$\max \left\{ U^{i*}\left(0, y_i^*, X_i\right); \hat{U}^{i*}\left(1, y_i^*, X_i\right) \right\}$$

where $U^{i*}\left(0, y_i^*, X_i\right)$ denotes the utility payoff to non participation defined in (11), and

$$\hat{U}^{i*}\left(1, y_i^*, X_i\right) = \max_{y} U^{i*}\left(1, y_i, X_i\right)$$

is the utility payoff of participation (i.e., $d_i = 1$), when the type of employment has been optimally selected by the household in the second stage. Since the labor force participation decision is binary, each household makes this choice by comparing the two payoffs.

Consider the net payoff of participation (i.e., the difference between the payoff of participating and the payoff of non participation). This net payoff, denoted D_i^*, writes as follows:

$$D_i^* := \hat{U}^{i*}\left(1, y_i^*, X_i\right) - U^{i*}\left(0, y_i^*, X_i\right).$$

To start our discussion on individual characteristics likely to raise women's employment probabilities, it is important to note that, singleness (i.e., $M_i = 0$) provides a woman with the incentive to work for pay, to support herself and any children she may have. Therefore, the discussion to follow, we focus exclusively on married women.

(i) Is Marriage a barrier to women's employment?

From expression (11), we obtain the net payoff of female employment as follows using (11), (16) and (20):

$$D_i^*|_{M_i=1} = \begin{cases} \bar{\beta}\log\left[\dfrac{\theta_i + \bar{A}E_i - \bar{c} - p_c S_i K_5^i}{\theta_i}\right] - \underline{\beta}\log\frac{(1+E_i)}{(1+\delta_0\ E_i)} & \text{if } y_i^* = 0 \\ \\ \bar{\beta}\log\left[\dfrac{\theta_i + \underline{A}E_i}{\theta_i}\right] - \underline{\beta}\log\frac{(1+E_i)}{(1+\delta_1\ E_i)} & \text{if } y_i^* = 1 \end{cases} \tag{26}$$

for all i such that $M_i = 1$. Since $\theta_i > 0$ for all co-headed households, it is clear that $D_i^*|_{M_i=1}$ is finite. Married women who participate in the paid labor force (i.e., $d_i = 1$) have individual characteristics X_i such that

$$D_i^*|_{M_i=1} > 0. \tag{27}$$

Those who do not participate (i.e., $d_i = 0$) have individual characteristics X_i such that

$$D_i^*|_{M_i=1} < 0. \tag{28}$$

Observe that for all y_i^*, the second term of expression (26) is strictly positive, whereas the sign of the first terms depends on the net economic contribution of a working married woman. This net contribution is $\bar{A}E_i - \bar{c} - p_c S_i K_5^i$ if she is wage-employed (i.e.,$y_i^* = 0$) and $\underline{A}E_i$, if self-employed. Therefore, to participate in the paid labor force, women must weigh the welfare gain from their net economic contribution against the welfare loss from a decrease in child quality (as measured by the second term of expression (26). A close inspection of this expression shows that whether marriage matters to women's participation in the paid labor force depends on a number of factors including (i) a woman's preferred type of employment (wage employment vs. self-employment), (ii) her post-marital living arrangements as determined by S_i, (iii) the size of her economic contribution conditional on chosen type of employment, and more importantly (iv) on her husband's earning θ_i. Indeed, it is clear from expression (26) that for all y_i^*:

$$\lim_{\theta_i \to 0} D_i^*|_{M_i=1} = \infty.$$

This implies that there a survival motive for married women's participation in the paid labor force. Hence the following result:

Proposition 2.3 *Let Assumptions 1-2 hold simultaneously. If θ_i is too small, then, the struggle survive*

will send married women into the paid labor force.

Proposition 2.3 states that poor household economic conditions are a driver of married women's participation in the paid labor force. It implies that marriage is not a barrier to female employment in poor countries. In other words, where marriage is a barrier to women's participation in the paid labor force, it must that the husband's income is sufficiently high, and the wife's economic contribution from paid work is lower than the induced loss in child quality.

(ii) Is motherhood a barrier to women's employment?

Again, for reasons stated above, to address this question we focus exclusively on married women (i.e., $M_i = 1$.

Assuming that (26) is differentiable in K_5^i, we distinguish between two cases.

First, consider a social setting where marriages leads to the emergence of extended family networks, i.e., $S_i = 0$. In this context, mothers with young children can rely on the help of relatives from the extended family network to provide nonparental childcare for free, if they wish to work for pay. In that case, expression (26) becomes independent of the number of children a woman has, K_5^i. This simply implies that motherhood does not matter to women's participation in a context where women have access to extended family networks to help provide free childcare to their young children. By contrast, in a context where families have a nuclear structure, so that extended family networks do not exist, i.e., $S_i = 1$, it can be shown through partial differentiation of expression (26) that $\Delta D_i^* \big|_{M_i=1}$ becomes strictly decreasing in K_5^i. This implies that motherhood act as a barrier to employment for women living in societies where families are nuclear. Hence the following result:

Proposition 2.4 *Let Assumptions 1-2 hold simultaneously. Then, motherhood has a negative effect on women's participation in the paid labor force only in a context where marriage leads to the formation of nuclear families. In a context where marriage leads to the emergence of extended family networks, motherhood does not matter to women's employment decisions.*

Proposition 2.4 implies that the social setting governs the importance of motherhood to female employment.

(iii) Does education promotes women's employment?

To address this issue, we take the partial derivative of (26) with respect to E_i:

$$\frac{\partial}{\partial E_i} D_i^*|_{M_i=1} = \begin{cases} \dfrac{\bar{A}}{\theta_i + \bar{A}E_i - \bar{c} - p_c S_i K_5^i} - \dfrac{\underline{\beta}\,(1-\delta_0)}{(1+E_i)\,(1+\delta_0\,E_i)} & \text{if } y_i^* = 0 \\ \\ \dfrac{\underline{A}}{\theta_i + \underline{A}E_i} - \dfrac{\underline{\beta}\,(1-\delta_1)}{(1+E_i)\,(1+\delta_1\,E_i)} & \text{if } y_i^* = 1 \end{cases}$$

As long as parental and nonparental caregiving are imperfect substitutes (i.e., in the sense that $\delta_0 < \delta_1 < 1$), the above derivative has an ambiguous sign. This is because female education generates a labor market return that trades off its return in terms of child development. On the one hand an educated woman who participates in the labor force brings home a higher income. But on the other hand, when nonparental caregiving imperfectly substitutes for parental caregiving, an educated woman who participates in the paid labor force adversely impact her young children's cognitive and socioemotional development. Because of these contrasting effects, education has an ambiguous effect on a woman's participation in the paid labor force.

To summarize, our model of how women's individual characteristics affect their employment outcomes leads to two clear predictions.

P.1. Education reduces women's self-employment probabilities.

P.2. The social setting governs the effect of motherhood on women's self-employment probabilities. In a social setting where marriage leads to the emergence of extended family networks, motherhood does not matter to women's self-employment.

P.3. Motherhood pushes women's facing poor household economic conditions into the paid labor force.

In what follows, we test these three predictions using micro-level data from Nigeria.

3 Background and Data

To investigate the effect of women's individual characteristics on their self-employment probabilities, we construct our main sample using data from Nigeria's 2018 *Demographic and Health Surveys* (DHS).

3.1 Women's Self-Employment in Nigeria

As the largest economy in sub-Saharan Africa, Nigeria has been creating more wage jobs for its growing population, with employment shifting slowly away from agriculture to wage work in private and public services (World Bank, 2015). But women remain marginalized from better-pay job opportunities. On the one hand, at 57%, Nigerian women's employment rates are among the highest in the (World Bank, 2015) and still rising. On the other hand, according to the same World Bank report, women in Nigeria are more likely than men to be in lower-earning occupations like farming and informal jobs. Indeed, women account for only 34% of wage employment but 60% of self-employment in the nonfarm sector in Nigeria. In particular, they make up the majority of the self-employed in the services sector. This fact raises the following question: what drives women's self-employment in Nigeria? Is it motherhood, as documented in developed countries (Connelly, 1992; Semykina, 2018) or some other factors?

At 5.4 children per woman in 2018, the fertility rate in Nigeria is among the highest in the world (Worl Bank, 2019). This fact suggests that women in this country still face enormous challenges to participation in the paid labor force, particularly during their most productive years, owing to childcare responsibilities. A high fertility rate thus may lend support to the hypothesis that, just like in developed countries, motherhood drives women's self-employment in Nigeria.

However, like many sub-Saharan African countries, Nigeria exhibits the practice of post-marital living arrangements whereby a married couple lives with or near the parents of at least one of the spouses after marriage. This type of post-marital arrangement leads to the creation of either multi-generation households—whereby children, parents, and grandparents live in the same home — or extended-family networks— whereby married couples live near the relatives of at least one of the spouses. In Nigeria, the proportion of multi-generation households— with both children under the age of 15 and family members aged 60 or higher— is 14%. In comparison, the corresponding figure is a mere 2% in the United States and Canada, respectively (United Nations, 2018). Post-marital living arrangements in Nigeria thus create a social setting where a significant proportion of married couples can rely on either grandparents or nearby relatives to help care for their young children when both parents work for pay. By contrast, in developed countries characterized by the prevalence of neolocal post-marital living arrange-

ments for married couples, the resulting nuclear family structure makes households dependent on costly professional childcare services (Pew Research Center, 2019). These facts suggest that motherhood may not be an essential factor for women's self-employment in Nigeria, when compared to, for example, Canada and the United States.

Indeed, roughly 44% of married women aged $20-24$ in 2017 married by the age of 18[6]. This fact suggests that early marriage is an essential factor of the persistently high fertility rate among Nigerian women (Corno et al., 2017). Further, evidence also shows that child marriage curtails female education (Field and Ambrus, 2008). The World Bank, for example, finds that, in Nigeria, girls are less likely to be in school than boys are (World Bank, 2015). Importantly, evidence shows that employment growth among Nigerian women has been driven by the labor force participation of women with the lowest levels of schooling (World Bank, 2015). These facts raise the critical issue of whether it is motherhood (as reflected in the high fertility rate) or lack of education that pushes Nigerian women into low-pay self-employment.

We use the 2018 round of Nigeria DHS to build a sample that will enable us to test the hypotheses discussed above.

3.2 Data

We construct our main analysis sample using Nigeria's 2018 DHS data. The DHS database contains information about respondents' socio-economic characteristics such as their education level, employment status, age group, marital status, birth history, household characteristics (e.g., household size, number of children, etc.), as well as their area and region of residence. Our sample contains 10 852 women aged between 15 and 49.

We use the women's questionnaire to construct two primary outcome variables: the employment and self-employment indicator functions. Employment is measured using an employment indicator function equal to 1 if a woman reported working in the last 12 months for pay, and 0 otherwise. Self-employment is measured using an indicator function equal to 1 if a working woman reported being self-employed for the last 12 months, and 0 if she said being wage-employed over the previous 12 months. We use the ILO definition of self-employment. The ILO

[6]See Girls Not Brides at https://www.girlsnotbrides.org/child-marriage/nigeria/

defines self-employed workers as persons of working age who are in one of the following categories: (a) employers; (b) own-account workers (or self-employed without hired employees); (c) contributing family workers and members of producers' cooperatives.[7] A wage-employed woman is someone who reported working for pay in a category other than one of the three self-employed types.

The questionnaire also contains several covariates of interest, including marital status, motherhood, and educational attainment. The marriage status indicator function equals 1 if a woman reported being married or cohabiting with a male partner for the last 12 months, and 0 if she said being single. Motherhood is a dummy that equals 1 if the female reported having at least one preschool child and 0 otherwise. The education variable is also an indicator function equal to 1 if a woman has at least a secondary education and 0 otherwise. We use secondary education as a threshold for considering a woman as educated. This conceptualization stems from the evidence showing that, unlike boys, girls often reap the economic and social benefits of education—e.g., increased lifetime earnings, dramatic decreases in fertility and mortality rates—only with participation in secondary education or higher (UNICEF et al., 2014).

Table (1) reports summary statistics in the sample for critical variables at both the individual and the regional levels. Wage-employed women have the highest level of education. Indeed, 54% of wage-employed women have at least a secondary education, while the corresponding figure for the self-employed is twice as low at 2.7%. The non-employed are the least educated: roughly 57% of them has no education. The corresponding figure for the self-employed is 46%, which is still very high when compared to the wage-employed, of which only 26% has no education. These summary statistics suggest that education positively influences women's wage employment.

On motherhood, women with no paid employment have a higher number of preschool children on average (0.73), followed by the self-employed (0.56), and the wage-employed with only 0.2 children on average. In other words, self-employed women have, on average, nearly three times as many preschool children as their wage-employed counterparts. This fact points to motherhood as having a positive influence on women's self-employment.

[7] see https://www.ilo.org/global/statistics-and-databases/statistics-overview-and-topics/status-in-employment/current-guidelines/lang–en/index.htm

Wage-employed women are also less likely to be married or cohabitating with a male partner (0.267) than self-employed women (0.536) or the non-employed (0.600). These statistics suggest that marriage may not be a barrier to women's employment in Nigeria.

Table (1) also presents the proportion of wage-employed, self-employed, and non-employed women in different age groups. The age distributions of self-employed and wage-employed women are almost identical. Nearly two-thirds (or 66.5%) of all self-employed women in our analysis sample reside in rural areas, compared to only 33.5% in urban areas. Wage employment is also predominantly rural, with nearly 60% of all wage-employed women residing in rural areas.

Table 1: *Summary Statistics*

	Wage-employed (n=5,156)		Self-employed (n=19,384)		Unemployed (n=14,227)	
VARIABLES	mean	sd	mean	sd	mean	sd
Sociodemographic characteristiques						
Married	0.781	0.414	0.877	0.329	0.796	0.403
No Education	0.265	0.442	0.460	0.498	0.556	0.497
Primary	0.198	0.398	0.270	0.444	0.122	0.328
At least secondary	0.537	0.499	0.270	0.444	0.321	0.467
Infertility or subfertility	0.100	0.301	0.119	0.324	0.142	0.349
Preschoolchildren	0.551	0.498	0.641	0.480	0.683	0.465
Children>5 at home	1.611	1.841	1.830	1.807	1.050	1. 683
Birth cohort						
15-19	0.0857	0.280	0.0383	0.192	0.225	0.418
20-24	0.128	0.334	0.107	0.309	0.223	0.416
25-29	0.158	0.364	0.162	0.369	0.175	0.380
30-34	0.133	0.340	0.144	0.351	0.116	0.320
35-39	0.128	0.334	0.141	0.348	0.0822	0.275
40-44	0.147	0.354	0.151	0.358	0.0769	0.266
45-49	0.220	0.415	0.258	0.438	0.101	0.301
Region						
North central	0.231	0.422	0.160	0.367	0.133	0.340
North east	0.158	0.364	0.172	0.377	0.287	0.452
North west	0.104	0.305	0.273	0.446	0.343	0.475
South east	0.114	0.318	0.102	0.303	0.0458	0.209
South south	0.218	0.413	0.135	0.342	0.0998	0.300
South west	0.176	0.381	0.157	0.364	0.0913	0.288
Area of residence						
rural	0.592	0.492	0.665	0.472	0.711	0.453

4 Empirical Strategy

In this section, we outline our empirical strategy for testing the main predictions of our theoretical model. In particular, the theory developed in the previous section suggests that a working woman becomes self-employed if her household's net utility payoff from picking this employment option is strictly positive. A linear approximation of the net utility payoff of self-employment can be written as a function of observed and unobserved factors, as follows:

$$y_{itr} = \mathbb{1}\left[\beta_0 + \gamma_t + \gamma_r + \beta_1 E_{itr} + \beta_2 K_5^{itr} + \beta_3 M_{itr} + \beta_4 R_{itr} + \epsilon_{itr} > 0\right], \tag{29}$$

where $y_{itr} \in \{0, 1\}$ denotes the self-employment status of a working woman i from birth cohort t residing in region r, and $\mathbb{1}[\cdot]$ is an indicator function equal to 1 if the inequality in parentheses holds true for woman i from birth cohort t residing in region r and 0 otherwise; E_{itr} is a dummy equal to 1 if a woman has at least a secondary education and 0 otherwise; K_5^{itr} is the motherhood status equals 1 if she has at least one preschool age child and 0 otherwise; M_{itr} her marital status, and R_{itr} the vector of other individual and household level controls. The term ϵ_{itr} is a mean-zero error term which captures the effect of unobserved factors that influence women's employment outcomes; γ_t and γ_r are respectively the birth cohort and the region fixed effects representing changes in self-employment risks related to her birth cohort t and to self-employment opportunities in her region of residence r. The main coefficient of interest is β_2. In settings where marriage leads to the creation of extended family networks, our theory predicts that $\beta_2 = 0$.

If all the covariates in the regression described in Equation (29) were exogenous, and all women in our empirical setting were employed, estimates obtained from (29) would be unbiased, thus yielding a correct identification of the causal effect of motherhood. However, the estimation of the self-employment equation in (29) raises two essential challenges. First, in our analysis sample, labor force participation among working-age women in our analysis sample is under two thirds, which increases the issue of whether selection into the labor force is nonrandom. Second, there is the potential endogeneity of our main covariates of interest, including education and motherhood.

4.1 Sample Selection Bias

Equation (29) is specified for the entire active population of women. However, the self- employment outcome is only observed among working women. Table (2) highlights differences in individuals characteristics between the sub-sample of working women and the sub-sample of those who are not gainfully employed. These differences do support the presence of sample selection bias in our empirical setting.

In our theory section, a selection rule is formulated by the net utility payoff of labor force participation whereby a woman participates in the paid labor force if and only if this net utility payoff is strictly positive. In this context, whether or not the self-employment outcome, y_{itr}, is observed depends on the following employment outcome:

$$d_{itr} = \mathbb{1}\left[\alpha_0 + \lambda_t + \lambda_r + \alpha_1 E_{itr} + \alpha_2 K_5^{ijr} + \alpha_3 M_{itr} + \alpha_4 X_{itr} + \eta_{itr} > 0\right] \tag{30}$$

where $\mathbb{1}[\cdot]$ is an indicator function equal to 1 if, for woman i of birth cohort t residing in region r, the (linearized) net utility payoff of participation in the paid labor force (i.e., the term in parentheses) is strictly positive, and zero otherwise;

η_{itr} is a zero-mean error term which captures the effect of unobserved factors that influence a woman's labor force participation decision [8]; λ_t and λ_r are respectively the birth cohort and the region's fixed effects for labor force participation, representing the changes in woman i's labor force participation related to her birth cohort t and to employment opportunities in her region of residence r.

A sample selection bias thus occurs when there exists a statistically significant correlation between unobservables in the errors terms in Equations (29) and (30). If sample selection were the only threat to the identification of the causal effects of motherhood on a woman's self-employment probability, estimation of the two-equation model represented by Equations (29) and (30) would produce consistent estimates of these effects.

[8] Heckman (1976, 1979) sample selection model assumes that the error terms ϵ_{itk} and η_{itr} are jointly normally distributed.

Table 2: *Differences in individual characteristics, between working, and non working women*

	Working Women		Not working women		T-test
Variable	Mean	SD	Mean	SD	p-value
Preschool age children	0.627	0.465	0.683	0.484	0.000
School age children	1.795	1.682	1.049	1.814	0.000
Married	0.862	0.403	0.796	0.345	0.000
At least secondary education	0.313	0.467	0.321	0.464	0.380
20-24	0.110	0.416	0.223	0.313	0.000
25-29	0.161	0.380	0.175	0.367	0.072
30-34	0.141	0.320	0.115	0.349	0.001
35-39	0.138	0.274	0.082	0.345	0.000
40-44	0.150	0.266	0.076	0.357	0.000
4549	0.252	0.301	0.101	0.434	0.000
Rural	0.653	0.453	0.711	0.476	0.000
North east	0.169	0.452	0.286	0.375	0.000
North west	0.245	0.474	0.343	0.430	0.000
South east	0.1042	0.209	0.045	0.305	0.000
South south	0.148	0.299	0.099	0.355	0.000
South west	0.160	0.288	0.0913	0.366	0.000

Significance threshold: *** p<0.01, ** p<0.05, * p<0.1

4.2 The Endogeneity of Motherhood

As argued in Semykina (2018), the literature on women's employment outcomes finds evidence of joint determination between fertility and women's labor force participation and labor supply, implying that a similar problem is likely to arise in the context of women's self-employment decisions. Likewise, the labor literature finds evidence of joint determination between an individual's education and employment decisions (Ahituv and Tienda, 2004), implying that this problem is also likely to arise in the context of women's self-employment decisions.

Following Agüero and Marks (2011), we use the union of two infertility/subfertility measures derived from survey respondents' self-reported infertility/subfertility. The first measure uses reported infertility as the reason for not currently using contraceptives (Infertile1). The second measure uses reported inability to have children as the reason for not achieving their desired number of children (Infertile 2). Our infertility indicator is:

$$Infertility = \max\{Infertility1, Infertility2\}.$$

Unlike instruments for fertility which target only women with children (such as twins at first birth (Bronars and Grogger, 1994) or the sex composition of the first two children (Angrist and

Evans, 1998; Cruces and Galiani, 2007)), this instrument enables both women with, and without, children to be included in the sample[9]

We estimate the following reduced form equation to measure the effect of infertility shocks on motherhood:

$$K_5^{ijr} = b_0 Infertility_{ij} + b_1(1 - Infertility_{ij}) + \psi_j + \psi_r + b_2 T_{ijr} + \nu_{ijr} \tag{31}$$

where K_5^{ijr} is the motherhood status of woman i of birth cohort j residing in region r, $Infertility_{ij}$ is her infertility status, ψ_j and ψ_r represent the fixed effects of the birth cohort and the region on women's fertility. T_{ijr} include control variables. The parameter β_0 captures the effect of being exposed to infertility on motherhood. According to Agüero and Marks (2011), this effect is expected to be negative and statistically significant.

Table 3: *Presence of other adults and children during the interview about sexual activity and women's self-reported infertility status*

Variable	In(sub-)fertile		Fertile		T-Test (b0=b1)
	b0	SD	b1	SD	p-value
Presence of husband adults during interview[10]	0.006	0.095	0.009	0.076	(0.314)
Presence of other male adults during interview	0.004	0.053	0.003	0.061	(0.574)
Presence of other female adults during interview	0.013	0.098	0.010	0.112	(0.283)

4.3 Estimation Method

To obtain consistent estimators of vectors α, β, **a**, and **b**, we estimate the following three-equation econometric model:

[9] Aguero and Marks (2008), Field and Ambrus (2008), Maccini and Yang (2009), Agüero and Marks (2011), and Shah and Steinberg (2017) argue that measurement errors do not present a significant problem for self-reported fertility shocks. Choosing this instrument thus eliminates the selection problem confronting the use of these other instruments.

An additional concern is that the presence of male family members at the time women answered the survey questionnaire may lead these women to misreport their fertility status. Our empirical results showed no evidence that the presence of other family members was correlated with responses to fertility questions (Results of this test are reported in Table (3)); see also Agüero and Marks (2011).

$$y_{itr} = \mathbb{1}\left[\beta_0 + \gamma_t + \gamma_r + \beta_1 E_{itr} + \beta_2 K_5^{itr} + \beta_3 M_{itr} + \beta_4 R_{itr} + \epsilon_{itr} > 0\right] \tag{32}$$

$$d_{itr} = \mathbb{1}\left[\alpha_0 + \lambda_t + \lambda_r + \alpha_1 E_{itr} + \alpha_2 K_5^{itr} + \alpha_3 M_{itr} + \alpha_4 X_{itr} + \eta_{itr} > 0\right] \tag{33}$$

$$K_5^{itr} = b_0 Infertility_{it} + b_1(1 - Infertility_{it}) + \psi_t + \psi_r + b_2 T_{itr} + \nu_{itr} \tag{34}$$

(32) is the self-employment (or main) equation. It describes the self-employment status, y_{itr}, of woman i of birth cohort t residing in region r, which is equal to 1 if self-employed and zero otherwise. Equation (33) is the employment (or selection) equation, describing women's selection into paid employment, such that $d_{itr} = 1$ if she participates in the paid labor force and zero otherwise. Importantly, as in Semykina (2018), we correct for the endogeneity of fertility in both the self-employment and the selection equations.

5 Empirical Results

In this section, we report and discuss the results of our estimation of the effects of women's characteristics on their self-employment probabilities. To assess the importance of correcting for nonrandom self-selection into the labor force and the endogeneity of motherhood, we follow Semykina (2018) in adopting a three-step approach. First, we estimate the self-employment equation separately (single-equation model). Second, we estimate a pair of two-equation models of women's self-employment in which we correct for nonrandom self-selection and the endogeneity of motherhood separately. Finally, we estimate a three-equation model that adjusts for these two problems jointly, as in Semykina (2018).

5.1 Accounting for Selection Bias and Endogeneity Separately

We start by presenting baseline estimation results, which correspond to step one and two described above.

Estimation results for steps one and two are reported in Table (4). When we do not correct for any potential identification problem (column 1, Table (4)), we find that motherhood has a

Table 4: *Correcting for selection and endogeneity of motherhood separately*

	One-Equation model	Two-Equation model (selection)		Two-Equation model (Endogeneity)	
	(1) y. eqn	(2) y. eqn	(3) s. eqn	(4) y. eqn	(5) K. eqn
Motherhood	0.070***	0.043***	0.049***	0.097**	
	(0.013)	(0.01)	(0.013)	(0.038)	
Education	-0.128***	-0.086***	0.001	-0.126***	
	(0.011)	(0.008)	(0.011)	(0.011)	
Infertility/subfertility					-0.256***
					(0.018)
School age children at home	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y	Y
Counterparts' unemployment rate	N	N	Y	N	N
Observations	7666	10,852	10,852	7666	7666

positive effect on women's self-employment probabilities. Having at least a preschool child increases the probability that a working woman is self-employed by 7 percentage points. This effect is statistically significant, and is consistent with findings in the literature, including Connelly (1992) and Semykina (2018). However, unlike Semykina (2018), we find that education has a negative and statistically significant effect on women's self-employment probabilities. Having at least a secondary education reduces the probability that a working woman is self-employed by nearly 13 percentage points (column 1, Table (4)). This result is significant because it signals that the social setting matters to women's self-employment probabilities. It implies that in the social setting of Nigeria—the empirical setting for this study— low educational attainment is a relatively more serious barrier to women's wage employment than motherhood is. The validity of the above results, however, hinges on the unrealistic assumptions that selection into the working women sample is random, and motherhood is exogenous. We explained further above why these two assumptions are likely to be rejected by the data, in which case estimates obtained from the one-equation model will be biased.

Indeed, when we correct only for nonrandom self-selection into the working women sample (see columns 2 and 3, Table (4)), results still show that motherhood has a positive and statistically significant effect on women's self-employment probabilities. However, the magnitude

of this effect is nearly 3 percentage points smaller than the effect obtained by estimating the one-equation probit regression (see columns 1 and 2, Table (4)). Education also maintains its negative impact on women's self-employment probabilities. However, the magnitude of this effect is more than 4 percentage points smaller than the effect obtained when estimating the one-equation probit regression model (see columns 1 and 2, Table (4)). Nevertheless, even in this case, low levels of education remain a relatively more critical driver of women's self-employment probabilities than motherhood.

In the selection equation (column 3, Table (4)), motherhood surprisingly has a positive effect on women's probabilities of participating in the labor force. Having at least one preschool child increases a woman's probability of employment by roughly 5 percentage points, and the result is highly significant. This result is quite the opposite of what Semykina (2018) finds in the socioeconomic context of the United States. Indeed, throughout sub-Saharan Africa, women are responsible for the food security of their families. They are expected to meet the basic survival needs of their children and other family members. They do so by working as laborers in family farms, by engaging in the farming of small animals and ruminants, in processing, particularly milk products and marketing farm products (Skoet and Stamoulis, 2006).

Furthermore, and in contrast to Semykina (2018), we find that education has no significant effect on women's employment probabilities. However, this result is valid only if sample selection is the only identification problem confronting our study. Next, when we estimate the self-employment equation jointly with the motherhood equation (column 4 and 5, Table (4)), Column 5 shows that infertility/subfertility is negatively correlated with motherhood, and this effect is highly significant. Infertile/subfertile women have 25.6 percent less chance of having a preschool child.

Column 4 shows that motherhood has a positive and statistically significant effect on women's self-employment. The presence of at least one preschool child increases a woman's self-employment probability by 9.7 percentage points, which is nearly 3 percentage points larger than the effect obtained by estimating the one-equation probit regression model. Once again, this effect is consistent with the existing literature and thus tends to support the conclusion that the social setting does not matter to the self-employment effect of motherhood. However, this conclusion is only valid if women's selection into the labor force is random.

Interestingly, education has a negative and highly significant effect on women's self-employment probabilities. Having at least a secondary education decreases a woman's probability of being self-employed by nearly 13 percentage points, which is similar in magnitude to the effect obtained by estimating the one-equation model. This result further contributes to the evidence that the social setting matters to women's self-employment probabilities. However, this result is obtained at the cost of ignoring the presence of sample selection bias reported above.

5.2 Accounting for Selection Bias and Endogeneity Jointly

The findings discussed above point to the fact that the positive effect of motherhood on women's self-employment probabilities extends to the social setting of Nigeria. Indeed, if consistent estimates could be obtained from at least one of the above specifications of our regression model, then this positive effect in Nigeria would amount to between 4 and 10 percentage points. However, we obtained the above findings by ignoring the fact that endogeneity and selection are both potentially present.

Here, we estimate a three-equation model to account for the facts mentioned above. In this three-equation model, we correct for selection bias and the endogeneity of the motherhood decision jointly. Tables (5) and (6) show the results of this estimation approach.

Table 5: *Correcting for selection and endogeneity of motherhood jointly*

	Three-equation model		
	(1)	(2)	(3)
	y.eqn	s. eqn	d. eqn
Motherhood	0.042	0.151***	
	(0.028)	(0.034)	
Education	-0.092***	0.049***	
	(0.009)	(0.012)	
Infertility			-0.266***
			(0.014)
School age children at home	Y	Y	Y
Birth Cohort FE	Y	Y	Y
Region FE	Y	Y	Y
Rural dummy	Y	Y	Y
Counterparts' unemployment rate	N	Y	N

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Table 6: *Estimated error correlation coefficients and test for endogeneity and selection*

Estimated error correlation coefficients	
	coefficients
$\rho_{21} = corr(\eta_{itr}, \nu_{itr})$	-0.216***
	(0.067)
$\rho_{31} = corr(\epsilon_{itr}, \eta_{itr})$	-0.883***
	(0.091)
$\rho_{32} = corr(\epsilon_{itr}, \nu_{itr})$	0.008
	(0.082)

Tests of endogeneity and sample selection		
	Statistic Chi(2)	p-value
Endogeneity (H0: $\rho_{21} = \rho_{32}$=0)	10.41	0.005
Sample Selection (H0: $\rho_{21} = \rho_{31} = 0$)	162.8	0.000

We find that motherhood has no statistically significant effect on women's self-employment (column 2, Table (5)). This result indicates that not correcting for selection and endogeneity jointly (Table (4)) leads to an overestimation of the effect of motherhood.

Correcting for sample selection bias and the endogeneity of motherhood jointly also shows that motherhood influences only women's employment probabilities. Having at least one preschool child increases a woman's probability of employment by roughly 15 percentage points (column 2, Table (5)). This surprising effect is highly significant. It is at odds with evidence from developed countries where motherhood reduces women's employment probabilities (Semykina, 2018). However, it is consistent with the fact that 63.3% of working-age women in our sample are employed, which is among the highest rates of female employment in the world[11]. Our theory explains this surprising result by the struggle to survive, which characterizes household economic conditions in many sub-Saharan African countries, including Nigeria. This struggle to survive forces both parents to work for pay, to cover their families' basic subsistence needs.

The effect of education on participation differs from its impact on the type of employment. Having at least a secondary education level increases a woman's probability of participating in the paid labor force by nearly 5 percentage points (column 2, Table 5). This result is consistent with Semykina (2018), who finds a positive effect of about $2-5$ percentage points in a devel-

[11]See World Bank data, available online at https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS

oped country's context. By contrast, having at least a secondary level of education reduces the probability that a working woman is self-employed by roughly 9 percentage points. This result differs from Semykina (2018), who finds that education has no effect on women's self-employment probabilities in a developed country's context. This surprising result stems from the fact that the return to education is higher in wage employment than in self-employment, as well as from the existence of post-marital living arrangements that create extended-family networks from which households can rely on to help with the care for young children.

6 Robustness checks

In this section we explore sources of heterogeneity in our main findings presented above. We examine several alternative specifications of our regression model, including stratification by area of residence and by marital status, as well as the use of an alternative estimation method.

6.1 Alternative Estimation Method: Two-Stage Residual Inclusion

In this subsection, we apply the two-stage residual inclusion method used by Schwiebert (2015) to control for endogeneity and selection bias when estimating the effect of motherhood on women self-employment. The results presented in Table (7) are similar to those obtained using the method of Semykina (2018). While the coefficient of preschool children is not significant, the effect of education is negative and statistically significant.

Results here confirm the robustness of our main findings. They increase our confidence that the social setting matters to the effect of motherhood on women's self-employment probabilities. In other words, the positive effect of motherhood documented in developed countries (Connelly, 1992; Semykina, 2018) is not universal. In a developing country's context, it depends on women's post-marital living arrangements for married couples.

6.2 Does the Area of Residence Matter?

Here we stratify our sample by area of residence, rural vs. urban. The results of this stratification are reported in Table 8. They show that motherhood does not matter to women's self-

Table 7: *Robustness check: Alternative estimation method*

	Second stage		*First- stage*
	(1)	(2)	(3)
	y.eqn	s. eqn	k. eqn
Motherhood	-0.028	0.154***	
	(0.028)	(0.034)	
Education	-0.085***	0.01	
	(0.007)	(0.011)	
First stage residual	0.075***	-0.119***	
	(0.029)	(0.036)	
Infertility/subfertility			-0.281***
			(0.01)
School age children at home	Y	Y	Y
Birth Cohort FE	Y	Y	Y
CUR	N	Y	N
Region FE	Y	Y	Y
Rural dummy	Y	Y	Y
Observations		10852	
Coefficient of correlation (second stage)			
corr(y eqn., d eqn.)			-0.956***
LR test of independence (rho = 0)			
Chi2 (1) statistic			26.76
Prob>chi2			0.0000

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

employment probabilities in rural areas as well as in the urban. Moreover, motherhood pushes women into the paid labor force in both rural and urban areas. Education reduces women's probabilities of self-employment in both in rural and urban areas. By contrast, unlike in urban areas were education increases the probability of employment, education in rural areas reduces this probability. A possible explanation of this result is that there tend to be very few employments opportunities for the educated in rural areas. We can, therefore, conclude that our main findings are robust to stratification by area of residence.

Table 8: *Robustness check, rural vs. urban*

	RURAL			URBAN		
	y.eqn	s. eqn	k. eqn	y.eqn	s. eqn	k. eqn
Motherhood	0.019	0.127***		0.104	0.245***	
	(0.026)	(0.045)		(0.084)	(0.071)	
Education	-0.053***	-0.069***		-0.147***	0.016	
	(0.011)	(0.016)		(0.024)	(0.019)	
Infertility/subfertility			-0.289***			-0.205***
			(0.018)			(0.024)
School age children at home	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y
Counterparts' unemployment rate	N	Y	N	N	Y	N
Observations	7272	7272	7272	3580	3580	3580

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

6.3 Does Marital Status Matter

In our main results, we find that motherhood does not matter to women's self-employment probabilities. We argue above that this finding shows that women's post-marital living arrangements create an extended family network on which women with young children can rely on to help care for their young children, should they decide to work for pay. However, these post-marital living arrangements apply only to married women.

In this subsection, we explore the implication of this fact for the effect of motherhood on the self-employment probabilities of single women. For this purpose, we stratify our analysis sample by marital status, distinguishing between married and unmarried women. The results of this stratification are reported in Table 9. For married women, we control for the husband/partner's education.

Table 9: *Robustness check: Single vs. married women*

	MARRIED WOMEN			SINGLE WOMEN		
	y.eqn	s. eqn	k. eqn	y.eqn	s. eqn	k. eqn
Motherhood	0.030	0.094*		0.172**	0.148	
	(0.052)	(0.05)		(0.068)	(0.20)	
Education	-0.10***	-0.015		-0.122	-0.034	
	(0.017)	(0.014)		(0.083)	(0.041)	
Infertility/subfertility			-0.252***			-0.161***
			(0.017)			(0.019)
Partner's Education	Y	Y	Y	N	N	N
School age children at home	Y	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y
Rural dummy	Y	Y	Y	Y	Y	Y
Counterparts' unemployment rate	N	Y	N	N	Y	N
Observations	9142	9142	9142	1710	1710	1710

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

We find that marital status does indeed matter for the effect of motherhood on women's self-employment probabilities. For married women, motherhood does not matter, which, as explained above, is consistent with a setting where extended family networks are an integral part of married couples' post-marital living arrangements.

For single women, by contrast, the picture is quite different. Having at least one preschool child increases the probability that a single woman engages in self-employment by roughly 17 percentage points. This result is significant at the 5% level. Our theory explains this result by the fact that female-headed households tend to function like nuclear families, where children are dependent on their parents for care.

We also find that motherhood has no significant effect on the labor force participation of single women. This result obtains because single women have to support themselves, which implies that they tend to work whether or not they have children.

Finally, as in the case of a developed country (Semykina, 2018), we find that education has no statistically significant effect on the self-employment probabilities of single women in a developing country context. This finding is likely to be the case if many single women have young children, which, in the absence of affordable nonparental care, restricts them to self-employment whether educated or not.

7 Conclusion

The literature finds that having young children can push women into self-employment to reconcile the demands of motherhood with professional ambitions. In this paper, we demonstrate that the social setting governs this prediction, causing it to fail in some countries. In particular, in social settings where self-employment offers few opportunities for decent pay, and marriage leads to post-marital living arrangements that support the emergence of extended family networks, women's self-employment may be driven by factors other than motherhood.

We formalize this using a theoretical model that explains how a woman's characteristics and the social setting in which she lives affect her labor force participation and self-employment decisions. Our model has three main features. First, we introduce women's post-marital living arrangements as practiced in their social setting as a critical determinant of the cost of caring for young children. Where marriage leads to the creation of multi-generation households or networks of extended family members living near one another, married women can rely on relatives to provide free childcare to their young children, should they decide to engage in paid work. By contrast, where marriage leads to the creation of nuclear families, women wishing to pursue wage employment must pay for professional childcare services for their young children. In this latter context, wage employment entails a motherhood penalty to working women. The second ingredient of the model is a woman's marital status. Post-marital living arrangements apply only to married women. Single women with young children tend to form a nuclear family, irrespective of the country in which they live. This makes them dependent on the availability of professional childcare providers if they choose to pursue wage employment opportunities. The final ingredient is the existence of (i) productivity premium for education in wage employment and (ii) a search cost for wage-paying jobs that is decreasing in a job seeker's level of education. The existence of a productivity premium for education implies that every worker in this context has an incentive to pursue wage employment instead of self-employment. This feature corresponds to a setting where self-employment takes place predominantly in the informal sector. The existence of a search cost for wage-paying jobs that decreases with the job seeker's level of education reflects a social setting where lack of schooling acts as a barrier to formal sector jobs where the bulk of wage employment takes place. We show that in societies

that feature the three ingredients mentioned above, having young children does not matter to women's self-employment if marriage leads to the emergence of extended family networks. Moreover, regardless of the custom of post-marital living arrangements, single women with young children are better off as self-employed.

We test these predictions using micro-level data from Nigeria, a significant proportion of households either live in multi-generation homes or closer to extended family members. Our empirical strategy accounts for the problems of nonrandom self-selection into the paid labor force and the endogeneity of fertility jointly.

We find no evidence that motherhood matters to women's self-employment in Nigeria. This finding is in contrast to results in the literature showing that this effect exists and is positive in a developed country's context (Connelly, 1992; Semykina, 2018). Second, we find that education has a negative and statistically significant effect on women's probabilities of self-employment. This finding also is in contrast to the existing literature showing that education has no statistically significant effect on this probability in a developed country's context (Semykina, 2018). We show that both these findings are robust to the use of an alternative estimation method, and the stratification of our sample by area of residence. These findings thus suggest that the reported positive effect of young children on women's self-employment is not universal. Furthermore, when we stratify our sample by women's marital status, we find that having young children matters to the self-employment of single women but not married women. We explain this finding by the fact that single women are most likely to rely on costly professional childcare providers because they tend to live on their own. This asymmetry in living arrangements between married and single women in a developing country's context may explain why young children have a positive effect on the self-employment probabilities of the latter but not the former.

Overall, our findings suggest that in a developing country's context, public policies aiming to reduce the gender gap in income by taking women out of low-pay self-employment must take into consideration women's marital status as well as their levels of education. Enhancing women's educational attainment may allow married women to turn their post-marital living arrangements into a free source of childcare services, enabling them to pursue better-pay wage employment. Taking women's marital status into consideration, by contrast, may allow policy-

makers to prioritize the childcare needs of single women, in the context of dwindling sources of government revenues.

References

Agüero, J., M. Marks, and N. Raykar (2012). The wage penalty for motherhood in developing countries. Technical report, Working Paper. Riverside, California: University of California, Riverside.

Aguero, J. M. and M. S. Marks (2008). Motherhood and female labor force participation: evidence from infertility shocks. *American Economic Review 98*(2), 500–504.

Agüero, J. M. and M. S. Marks (2011). Motherhood and female labor supply in the developing world evidence from infertility shocks. *Journal of Human Resources 46*(4), 800–826.

Ahituv, A. and M. Tienda (2004). Employment, motherhood, and school continuation decisions of young white, black, and hispanic women. *Journal of Labor Economics 22*(1), 115–158.

Angrist, J. D. and W. N. Evans (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review 88*(3), 450–477.

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of political Economy 116*(4), 709–745.

Baker, M. and K. Milligan (2015). Maternity leave and children's cognitive and behavioral development. *Journal of Population Economics 28*(2), 373–391.

Basker, E. (2012). Raising the barcode scanner: Technology and productivity in the retail sector. *American Economic Journal: Applied Economics 4*(3), 1–27.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *Journal of political Economy 81*(2, Part 2), S279–S288.

Becker, G. S. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of political Economy 84*(4, Part 2), S143–S162.

Bertrand, M., P. Cortés, C. Olivetti, and J. Pan (2016). Social norms, labor market opportunities, and the marriage gap for skilled women. Technical report, National Bureau of Economic Research.

Bianchi, S. M. (2000). Maternal employment and time with children: Dramatic change or surprising continuity? *Demography 37*(4), 401–414.

Boserup, E., S. F. Tan, and C. Toulmin (2013). *Woman's role in economic development*. Routledge.

Bronars, S. G. and J. Grogger (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *The American Economic Review*, 1141–1156.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). A flying start? maternity leave benefits and long-run outcomes of children. *Journal of Political Economy 123*(2), 365–412.

Casarico, A. and A. Sommacal (2018). Taxation and parental time allocation under different assumptions on altruism. *International Tax and Public Finance 25*(1), 140–165.

Conference, I. L. (2013). *Transitioning from the Informal to the Formal Economy*. International Labour Office.

Connelly, R. (1992). Self-employment and providing child care. *Demography 29*(1), 17–29.

Corno, L., N. Hildebrandt, and A. Voena (2017). Age of marriage, weather shocks, and the direction of marriage payments. Technical report, National Bureau of Economic Research.

Cruces, G. and S. Galiani (2007). Fertility and female labor supply in latin america: New causal evidence. *Labour Economics 14*(3), 565–573.

Currie, J. and E. Moretti (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics 118*(4), 1495–1532.

Danzer, N. and V. Lavy (2018). Paid parental leave and children's schooling outcomes. *The Economic Journal 128*(608), 81–117.

De La Croix, D. and M. Doepke (2003). Inequality and growth: why differential fertility matters. *American Economic Review 93*(4), 1091–1113.

De la Croix, D. and M. Doepke (2004). Public versus private education when differential fertility matters. *Journal of Development Economics 73*(2), 607–629.

Del Boca, D., C. Flinn, and M. Wiswall (2013). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Drange, N. and T. Havnes (2014). Early and bright? child care for toddlers and early cognitive skills. In *Manuscript presented at the EALE Conference in Ljubljana*.

Drange, N. and T. Havnes (2019). Early childcare and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics 37*(2), 581–620.

Dustmann, C. and U. Schönberg (2012). Expansions in maternity leave coverage and children's long-term outcomes. *American Economic Journal: Applied Economics 4*(3), 190–224.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Field, E. and A. Ambrus (2008). Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy 116*(5), 881–930.

Fort, M., A. Ichino, and G. Zanella (2017). The cognitive cost of daycare 0-2 for children in advantaged families. In *Technical Report*. Working Paper.

Garcia-Mainar, I. and V. M. Montuenga-Gomez (2005). Education returns of wage earners and self-employed workers: Portugal vs. spain. *Economics of Education Review 24*(2), 161–170.

Gupta, N. D. and M. Simonsen (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics 94*(1-2), 30–43.

Heath, R. (2017). Fertility at work: Children and women's labor market outcomes in urban ghana. *Journal of Development Economics 126*, 190–214.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement, Volume 5, number 4*, pp. 475–492. NBER.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Herbst, C. M. (2013). The impact of non-parental child care on child development: Evidence from the summer participation "dip". *Journal of Public Economics 105*, 86–105.

Kikuchi, T. (2007). Network externalities and comparative advantage. *Bulletin of Economic Research 59*(4), 327–337.

Lim, K. (2018). Do american mothers use self-employment as a flexible work alternative? *Review of Economics of the Household*, 1–38.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–26.

Mercer, J. (2006). *Understanding attachment: Parenting, child care, and emotional development*. Greenwood Publishing Group.

Moock, P. R., P. Musgrove, and M. Stelcner (1989). *Education and earnings in Peru's informal nonfarm family enterprises*, Volume 236. World Bank Publications.

Noboa-Hidalgo, G. E. and S. S. Urzua (2012). The effects of participation in public child care centers: Evidence from chile. *Journal of Human Capital 6*(1), 1–34.

Pew Research Center (2019). Religion and living arrangements around the world. *retrieved online at https://www.pewforum.org/2019/12/12/religion-and-living-arrangements-around-the-world/*.

Rasmussen, A. W. (2010). Increasing the length of parents' birth-related leave: The effect on children's long-term educational outcomes. *Labour Economics 17*(1), 91–100.

Roche, K. (2014). Female self-employment in the united states: an update to 2012. *Monthly Lab. Rev. 137*, 1.

Schwiebert, J. (2015). Estimation and interpretation of a heckman selection model with endogenous covariates. *Empirical Economics 49*(2), 675–703.

Semykina, A. (2018). Self-employment among women: Do children matter more than we previously thought? *Journal of Applied Econometrics 33*(3), 416–434.

Shah, M. and B. M. Steinberg (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy 125*(2), 527–561.

Skoet, J. and K. Stamoulis (2006). *The state of Food Insecurity in the World*. Food and Agriculture Organization of the United Nations.

Tertilt, M. (2006). Polygyny, women's rights, and development. *Journal of the European Economic Association 4*(2-3), 523–530.

UNICEF et al. (2014). Accelerating secondary education for girls: Focusing on access and retention.

United Nations (2018). Composition around the world 2017. *United Nations' Data Booklet. URL: https://www. un. org/en/development/desa/population/publications/pdf/ageing/household_size_and_composition_around_the_world_2017_data_booklet. pdf. Title from the screen 18*.

Wellington, A. J. (2006). Self-employment: the new solution for balancing family and career? *Labour Economics 13*(3), 357–386.

Worl Bank (2019). Fertility rate, total (births per woman). *World Bank, Country data, https://data.worldbank.org/indicator/SP.DYN.TFRT.IN.*.

World Bank (2015). More, and more productive jobs for nigeria: A profile of work and workers. Technical report, International Bank for Reconstruction and Development / The World Bank.