DISCUSSION PAPER SERIES

No. 3026

WEALTH DISTRIBUTION AND SELF-EMPLOYMENT IN A DEVELOPING ECONOMY

Alice Mesnard and Martin Ravallion

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3026.asp**

ISSN 0265-8003

WEALTH DISTRIBUTION AND SELF-EMPLOYMENT IN A DEVELOPING ECONOMY

Alice Mesnard, Université des Sciences Sociales de Toulouse and CEPR
Martin Ravallion, World Bank

Discussion Paper No. 3026
October 2001

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3026

October 2001

ABSTRACT

Wealth Distribution and Self-Employment in a Developing Economy*

The extent of entrepreneurial activity in an economy with poorly developed capital markets depends on the distribution of wealth, though in potentially complex ways. A non-parametric model of the wealth effect on self-employment is estimated using micro data on the occupational choices of return migrants in Tunisia. Controls for heterogeneity are included, and tests are made for selection bias and separability between wealth and the controls. There is no sign of increasing returns at low wealth, suggesting generally low start-up costs in this setting. The aggregate self-employment rate is an increasing function of aggregate wealth, but a decreasing function of wealth inequality, though even a substantial redistribution of wealth would have only a small impact.

JEL Classification: D31 and M13
Keywords: borrowing constraints, migration, self-employment and wealth inequality

Alice Mesnard
ARQADE
Université des Sciences Sociales
de Toulouse 1
21 Allée de Brienne
31 000 Toulouse
FRANCE
Tel: (33 5) 6112 8529
Fax: (33 5) 6112 8538
Email: alice.mesnard@univ-tlse1.fr

Martin Ravallion
Development Research Group
World Bank, Room MC3-581
1818 H Street NW
Washington DC 20433
USA
Tel: (1 202) 473 6859
Fax: (1 202) 522 1153
Email: mravallion@worldbank.org

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=152406

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=138943

* Work on this Paper began while Ravallion was visiting the University of Toulouse, whose hospitality is gratefully acknowledged. These are the views of the authors, and need not reflect those of their employers including the World Bank. For their comments on this paper, the authors are grateful to Bruno Biais, Pierre Dubois, Esther Duflo, Emanuela Galasso, Vassilis Hajivassiliou, Peter Lanjouw and Dominique van de Walle.

Submitted 19 September 2001

1. Introduction

Self-employment is an important element of the private economy in developing countries. For example, it is estimated that 34% of the nonagricultural labor force of North Africa in 1990 was self-employed (up from 12% in 1970) (United Nations, 2000).[2] The bulk of this self-employment is found within what is often referred to as the "informal sector" and one of the stylized facts about the sector is that entry costs tend to be much lower than for the formal sector (Thomas, 1992, Chapter 4). This is in marked contrast to developed countries where self-employment is more often part of the dominant formal segment of the economy for which entry costs are clearly non-negligible.

In attempting to understand why some people become self-employed and others do not, a number of studies (for both developed and developing countries) have emphasised the importance of current wealth.[3] Capital market failures have been seen as the most likely reason for positive wealth effects on the probability of self-employment. By implication, the level of aggregate wealth matters to aggregate self-employment. But does the distribution of aggregate wealth also matter? This question is of interest as a clue to understanding the diverse development paths taken by countries starting at similar average incomes. Banerjee and Newman (1993) provide a theoretical model in which wealth inequality influences occupational structure, given capital market failures. (In a similar vein see Aghion and Bolton, 1997.)

The literature has offered almost no evidence that the distribution of wealth matters to employment structure. In the only exception we know of, Lindh and Ohlsson (1998) argue that

---

2 Historically, family businesses were also a prominent feature in the early stages of the economic development of today's developed countries (Bhattacharya and Ravikumar, 2001).

3 See Evans and Jovanovic (1989), Evans and Leighton (1989), Holtz-Eakin, Jouflaian and Rosen (1994), Van Praag and Van Ophem (1995), Magnac and Robin (1996), Lindh and Ohlsson (1996), Blanchflower and Oswald (1998), Dunn and Holtz-Eakin (2000), Paulson and Townsend (2000).

falling wealth inequality over time in Sweden has attenuated entry into self-employment. They interpret this as the combined effect of borrowing constraints and start-up costs (echoing the theoretical model of wealth dynamics in Aghion and Bolton, 1997). By this view, greater wealth equality means that fewer potential entrepreneurs are able to finance the required start-up capital.

This paper examines the influence of wealth inequality on aggregate self-employment in a developing economy. We focus on business starts (rather than business longevity or profitability) because this provides a convenient window on aggregate business activity.[4] Our theoretical model aims to capture the inherent ambiguity in the effect of wealth inequality on self-employment. In explaining occupational choice, we follow Evans and Jovanovic (1989), Banerjee and Newman (1993) and others in emphasizing the role played by capital market imperfections rather than differences in preferences.[5] Borrowing constraints entail that the curvature of the own-wealth effect on the probability of becoming self-employed depends heavily on the shape of the production function. This holds when exogenous borrowing constraints take the common (homogeneous linear) form found in the literature following Evans and Jovanovic (1989). We also show that the same feature can be obtained with a more general model in which borrowing constraints emerge from endogenous credit limit, given moral hazard.

An implication of our model is that higher wealth inequality reduces the aggregate rate of self-employment when start-up costs are small. However, the effect goes in the opposite direction with sufficiently high start-up costs — generating non-convexities in the set of employment opportunities at low levels of wealth. Since start-up requirements tend to be smaller

4 One can conjecture that credit constraints also impede business development and longevity, and there is some supportive evidence (Bates, 1990; Levy, 1993). However, we do not know of any evidence that wealth inequality matters to aggregate profitability and/or longevity.

5 Preference-based approaches have instead emphasized differences in attitudes to risk, as in Kihlstrom and Laffont (1979) and Kanbur (1979).

in developing countries than developed ones, our theoretical model suggests that wealth inequality can have opposite effects on employment structure between rich and poor countries.

Motivated by this model, we look for nonlinear wealth effects on self-employment using micro data on return migrants in a developing country. Our setting and methods differ from past work in a number of respects.[6] Studying return migrants offers the hope of avoiding a potentially serious problem in past work on self-employment using micro data, namely that wealth immediately prior to taking up self employment is typically unobserved. Wealth data are normally collected at the survey date — after the occupational choice has been made. A concern about past work is that wealth data collected this way may be endogenous to occupational choices.[7] In contrast, migrants have returned to their home country with diverse amounts of wealth, accumulated from savings while abroad, and they face the choice of what to do on their return. There is a natural dynamic sequence. The fact of being predetermined does not of course guarantee exogeneity; possibly some latent propensity for self-employment prompted greater wealth accumulation while abroad.[8] However, the sequencing clearly offers a better prospect for identifying wealth effects than in regular data sets in which one jointly observes wealth and occupational choice. By asking how much they accumulated while abroad, and seeing which

---

6 In the only previous empirical work on the effect of inequality on self-employment, Lindh and Ohlsson (1998) used time series data on aggregate self-employment and the parameters of the distribution of wealth in Sweden, for which they had 16 observations (spanning 1920-92). The wealth distribution was assumed to be exogenous and no controls were used for other possible determinants of the long-run evolution of employment structure of the economy. However, they did find that the correlation with wealth distribution was robust to de-trending the data.

7 Dunn and Holtz-Eakin (2000) comment on this concern about past empirical work. Their solution is to study the transition from wage employment to self-employment (rather than self-employment per se) and to combine this with data on wealth and other characteristics at a time prior to this transition. Studying return migrants and their wealth at the date of return serves a similar function.

8 Mesnard (1999) develops this argument further in a formal model.

return migrants started a new business, we test for nonlinearity in the wealth effect, consistent with the claim that wealth distribution matters to aggregate self-employment.[9]

We depart from past empirical work on the determinants of self-employment in that we do not parameterize the relationship with wealth. As our theoretical model makes clear, there is no obvious reason why the wealth effect would be linear, or indeed have any specific parametric form. If there are fixed start-up costs, with diminishing returns setting in after, and the liquidity constraint takes the form of a collateral requirement for borrowing, then one would expect the effect to be highly nonlinear, at least amongst observationally identical people. So we treat the wealth effect in a flexible way, using non-parametric methods. We also allow for heterogeneity in the form of linear controls. Thus we draw on recent advances in the estimation of partial linear models (Yatchew, 1997, 1998).

The following section discusses the partial equilibrium effect of wealth on business start-ups in general terms, while section 3 outlines our econometric model. Section 4 describes the data and gives some descriptive results. The estimation results are presented in section 5, while section 6 provides some tests of robustness. Section 7 examines their implications.

2. Wealth effects on self-employment in theory

A given amount of accumulated savings $W_i\ (\geq 0)$ allows person *i* to borrow up to some amount of start-up capital $K_i \geq W_i$. We initially follow Evans and Jovanovic (1989) in assuming that one can borrow up to some fixed multiple of wealth, i.e.,

$$K_i / W_i = k \geq 1 \text{ for all } i \text{ and } W_i \in [W^{\min}, W^{\max}] \tag{1}$$

9 We know of two studies of occupational choices of return migrants in developing countries, both of which found that higher accumulated savings while abroad are associated with higher probabilities of entering self-employment relative to wage labor (Mesnard, 1999; Ilahi, 1999). We know of no previous attempt to examine the nonlinearity in this wealth effect on occupational choice.

This characterizes what one would expect to find if poor entrepreneurs have little access to credit, but borrowing becomes easier (though still credit-constrained) at higher levels of wealth. The precise form of equation (1) is *ad hoc*; later we consider a model with an endogenous borrowing constraint, though this does not change our main message.

In characterizing self-employed production, it is obvious that many businesses require a positive minimum level of capital to start operations. However, the informal sector activities one finds in developing countries often appear to have modest start-up requirements, though possibly still entailing non-negligible costs for poor people. The possibility of start-up costs creating a non-convexity in the set of employment opportunities at low wealth will be one of the things we shall look for in our empirical work, but for now we remain eclectic on this point.

The output from the new business is $F(K_i)$ where *F* is smoothly increasing throughout, and strictly concave above some capital stock, $K^* \geq 0$. To allow increasing returns at low wealth, we assume that *F* is convex for $K < K^*$ (when $K^* > 0$). There is some fixed own-labor requirement that is subsumed in the function *F*. The (risk-free) opportunity cost of capital is *r* while for own labor it is ω_i. The latter is unobserved to us (though we will allow for plausible but partial covariates in the empirical work). For the present purpose, we treat ω_i as a random variable with continuous twice-differentiable distribution function Ψ. (Note that ω_i also includes any idiosyncratic additive differences in output at given capital stock.)

If $F(K_i) - rK_i$ exceeds ω_i then (and only then) will person *i* become self-employed. The probability of becoming self-employed (S_i) is then given by:

$$S_i \equiv S(W_i) = \Psi[F(K_i) - rK_i] \quad (2)$$

with slope:

$$S'(W_i) = \Psi'(.)k[F'(K_i) - r] \quad (3)$$

This vanishes in the special case in which the worker is not liquidity constrained, and can employ as much capital as desired ($F'(K_i) = r$). In the liquidity-constrained case, $S'(W_i) > 0$.

The aggregate number of new business start-ups in this economy is $\sum S(W_i)$ which can be thought of as a function of the distribution of wealth, $(W_1, W_2, ..., W_n)$. Consider two distributions A and B, with the same mean. From well-known properties of concave functions, if distribution A has higher inequality than B — in that A can be obtained from B by a set of transfers in which the donor has lower wealth than the recipient — then A will generate a higher (lower) aggregate number of new business starts than B if the function *S* is convex (concave).

So we need to consider the curvature of *S*. In the special case in which ω_i has a uniform distribution ($\Psi'(.)$ constant), it is plain that *S* will have the same curvature as *F*:

$$S''(W_i) = \Psi'(.)k^2 F''(K_i) \quad (4)$$

If diminishing returns set in from low wealth ($K^* = 0$), then *S* will be concave throughout. For $K^* > 0$ we obtain Figure 1, with *S* convex at low wealth (below K^*/k) and concave above.

If diminishing returns set in from the outset ($K^* = 0$) then higher inequality will unambiguously reduce the rate of self-employment. However, the effect is ambiguous for $K^* > 0$, in that the outcome will depend on exactly how the change in distribution is achieved. Inequality amongst those above the point of inflexion in Figure 1 will reduce self-employment; inequality amongst those with less than this value will increase it. Small mean-preserving redistributions from anyone above the point of inflexion in Figure 1 to anyone below it could either increase or decrease the aggregate number of business starts in a fixed population. Nothing can be said about redistributions from those above this point to those below it.

The same basic picture emerges under weaker assumptions, though one can also derive more complex forms of nonlinearity, and (hence) more complex forms of distribution dependence. Let us first relax the assumption that access to capital is proportional to current wealth. Instead of (1), assume that $K_i = K(W_i)$ where $K(W_i) \geq W_i$ with $K'(W_i) > 0$ for all W and a second derivative that can be either positive or negative. Then:

$$\begin{aligned} S''(W_i) = \Psi'(.)[K'(W_i)^2 F''(K_i) + (F'(K_i) - r)K''(W_i)] \\ + \Psi''(.)(F'(K_i) - r)^2 K'(W_i)^2 \end{aligned} \tag{5}$$

It is evident from (5) that the curvature of S is now ambiguous even with $K^* = 0$. There is a concave effect coming from diminishing returns ($K'(W_i)^2 F''(K_i) < 0$). But there is an effect of unknown curvature from the borrowing constraint $(F'(K_i) - r)K''(W_i)$) and a third term, also of unknown sign, coming from any non-uniformity in the distribution of labor cost.

To throw more light on the curvature of the borrowing constraint, let us now see how the function $K(W_i)$ can be derived endogenously. We follow Banerjee and Newman (1993) in identifying $K_i - W_i$ as the maximum that will be leant to someone with collateral W_i, taking account of the borrower's expected gains from default. (Though the details of our model differ somewhat from Banerjee and Newman.) Defaulters are apprehended with probability $\lambda > 0$. If the borrower defaults and is not caught, then we assume that she obtains the full output from the enterprise, $F(K_i)$. However, if she defaults and is caught then the lender confiscates all output. So the expected payoff from defaulting is $(1-\lambda)F(K_i)$. The payoff from not defaulting is the output from the enterprise plus the refunded collateral less the loan repayment with interest; so the payoff is $F(K_i) + W_i - (1+r)K_i$. On equating this with the expected payoff from defaulting, $K(W_i)$ solves $W_i = (1+r)K_i - \lambda F(K_i)$. On differentiating w.r.t. W_i, it is readily verified that

$K'(W_i) = [1 + r - \lambda F'(K_i)]^{-1}$ (for $K'(W) > 0$ we require that the default probability is bounded above by $(1+r)/F'(K_i)$). Substituting $K''(W_i) = \lambda F''(K_i) K'(W_i)^3$ into (5) we find that:

$$S''(W_i) = \Psi'(.)K'(W_i)^3 F''(K_i)[1 + r(1-\lambda)] + \Psi''(.)(F'(K_i) - r)^2 K'(W_i)^2 \qquad (6)$$

In the case of a uniform distribution of labor cost, the curvature of the function *S* is again determined by the curvature of the production function.

To allow a non-uniform distribution of labor cost, suppose that the distribution is unimodal. The density is rising at low wealth ($\Psi''[F(K(W^{\min})) - rK(W^{\min})] > 0$) and falling at high wealth ($\Psi''[F(K(W^{\max})) - rK(W^{\max})] < 0$). This is not sufficient to assure that the curvature of the function *S* is the same as that of the production function, and stronger assumptions are required to assure a unique point of inflexion in *S*. However, it is evident that the function will again be convex at low wealth and concave at high wealth as in Figure 1. The ambiguity in the distribution dependence of aggregate self-employment will remain even if diminishing returns to capital set in immediately ($K^* = 0$, so $F''(K_i) < 0$ for all *K* in (6)).

3. An empirical model

The above discussion points to the need to represent the wealth effect in a flexible way. Non-parametric regression methods are an obvious choice. However, we also want to allow for heterogeneity in non-wealth characteristics. This could arise from differences in output at given capital (such as due to the availability of family labor to help with the business) or differences in the opportunity cost of labor. To make the estimation tractable, given the non-parametric regression function, we assume a linear probability model in the model with controls. With linear controls, our method entails estimating partial linear regressions, in which the sub-function for the wealth effect is nonparametric.

Linearity in the control variables has the well-known shortcoming that (unlike a probit or logit model) there is no guarantee that the predicted values for *S* will lie within the [0,1] interval. We will check if our estimates satisfy this condition. However, even if satisfied at all data points, that check cannot be considered conclusive since the parameter estimates on which it is based will be inconsistent if the underlying probability model is misspecified. This is of greater concern if the true probabilities are close to the extremes where violations of the assumed linearity can be expected to be less negligible.[10] We will compare our results for a partial linear model with a parametric model using a (parametric) nonlinear probability function.

We write the probability of becoming self-employed as some unknown function of *W*, as well as control variables *X*:

$$S_i = \phi(W_i) + X_i\pi + \nu_i \quad (i = 1,...,n) \tag{7}$$

in which the zero-mean innovation error has variance σ_ν^2. All that we assume about the function ϕ is that it is smooth and single valued; in particular, its first derivatives are bounded by constants, $c \geq |\Delta\phi(W_i)|/|\Delta W_i|$. The function need not be monotonic, or take any parametric form. Equation (7) imposes separability between *W* and *X*. Later we offer a specification test of this property, by seeing if there is any correlation between the estimated residuals of (7) and the interaction effects between the *X*'s and the slopes of the ϕ function. (Such an interaction effect could arise if credit markets work better for some people than others. For example, wealth may be less of a constraint on starting a business for well-educated entrepreneurs.)

---

10 Matzkin (1992) shows how if certain *a priori* restrictions are placed on the non-parametric regression function then consistent estimation is possible under weaker assumptions about the error distribution than we have made here. As formulated, our theoretical model does not lend itself to the restrictions needed for applying the Matzkin method. However, it may be possible in future work to find economically interesting specializations of our model that facilitate application of this method.

To estimate the model of business starts in (7) we draw on the literature on partial linear models. Following Yatchew (1997), we order all observations in terms of their values of W_i and take differences between the data for successive ranked observations, giving the regression:

$$\Delta S_i = \Delta\phi(W_i) + \Delta X_i \pi + \Delta \nu_i \tag{8}$$

where ΔX_i is the difference between the value of *X* for the *i*'th observation and that for *i*-1 when ranked in ascending order of *W*. Under our assumption about the function ϕ, the first term on the RHS vanishes as *n* goes to infinity ($\text{plim}[\phi(W_i) - \phi(W_{i-1})] = 0$). So we estimate the following parametric regression by least squares:

$$\Delta S_i = \Delta X_i \pi + \Delta \nu_i \tag{9}$$

We then estimate the non-parametric regression:

$$S_i - X_i \hat{\pi} = \phi(W_i) + \nu_i \tag{10}$$

Higher-order differencing allows efficiency gains (Yatchew, 1997). We write (9) as:

$$\sum_{j=0}^{m} d_j S_{i-j} = (\sum_{j=0}^{m} d_j X_{i-j})\pi + \sum_{j=0}^{m} d_j \nu_{i-j} \tag{11}$$

where $\sum d_j = 0$ (which allows us to drop the non-parametric effect from equation 11) and the normalization condition $\sum d_j^2 = 1$ (which assures that the transformed residuals have variance σ_ν^2). Hall et al., (1990) provide the optimal weights up to *m*=10.

4. Data and descriptive statistics

Our data are from a survey that was done in 1989 by the Tunisians Settled Abroad Office in the Foreign Affairs Ministry, with the collaboration of the Arabic League.[11] The survey was

[11] See Mesnard (1999, pp. 205-211) for a more detailed description of the survey.

conducted in all geographical areas of Tunisia (both rural and urban areas). Return migrants are defined as workers who have worked abroad at least once during 1974-86 and have returned to live in Tunisia before the survey date. This group was over-sampled to get more observations.

The non-migrant sample suggests only a weak correlation between wealth and self-employment. From the survey we can measure the wealth accumulated up to 1986. The mean wealth of those who became self-employed between 1974 and 1986 is 689 dinars, while it is 454 for those who did not. However, the difference is not statistically significant (a t-test gives 1.62). If we restrict the sample to the workers who were not self-employed in 1974, the difference between the mean wealth accumulated up to 1986 by self-employed (equal to 689 dinars) and by salaried workers (397 dinars) becomes significant (with a t-test of 2.02).

However, it could be hard to detect credit-constraints in the non-migrant sample. We know whether the respondent is self-employed at the date of interview. And we know wealth accumulation up to that date. However, there are obvious concerns about the endogeneity of wealth to self-employment in the sample of non-migrants (similarly to the concerns about past empirical work on the determinants of self-employment, as discussed in the introduction).

We focus instead on the sample of returned migrants. Returning from a long period overseas makes a natural break in work history. For such migrants we can measure the probability of self-employment after returning to the home country. And we can identify how much money they brought back from their period overseas, which is pre-determined at the time they make their decision about what work to do on return. This group is also more homogeneous than the population at large, which should also make it easier to detect any relationship between wealth and occupational choice.

The return migrants can be thought of as a sample selected from a complete set of migrants, not all of whom returned. The sample selection process is unknown and we have no data on those who did not return. It is plausible that the probability of returning depends positively on accumulated wealth. Then the wealth effects we see in the data on return migrants also include an unknown selection effect, as well as liquidity constraints.

However, this is not a serious concern for the problem at hand. We want to know how important wealth distribution within the home country is to the level of business activity in that country. Since we want to condition on residence in the home country, we need not be concerned about the possibility that return migrants are untypical of migrants that did not return.

Another possible selection bias is more worrying. Return migrants may not be typical of the workforce as a whole in the country they return to. This we can test for, using data from the sample of non-migrants. We discuss the details later.

Our main sample covers 1224 male returned migrants who reported that they intend to stay indefinitely in Tunisia.[12] The survey obtained general information about their migration history (number of migrations, dates, locations, return motives, duration, employment) and their working and living conditions during their last migration. To identify new business starts in the data we build a dummy variable equal to one if a worker is self-employed after return and was not self-employed before migrating and equal to zero otherwise. The new businesses tended to be started up quite quickly after return (45% began within two months). There is comprehensive information on the assets accumulated during their migration. The survey obtained data on a number of obvious control variables, including age and education. Whether one takes up self-employment is also likely to be correlated with where one lives, as an influence on proximity to

[12] We dropped 102 migrants who are temporarily visiting Tunisia for vacation, as well as 12 women.

markets. However, endogeneity concerns speak against controlling for current location in this context (given that return migrants can in principle choose where to return to). We do include, instead, controls for place of birth.

Table 1 provides summary statistics. Column (1) gives data on the sub-sample who had not been self-employed prior to migrating but took up self-employment on return. Column (2) gives data for those who had not been self-employed, and did not take this up on return. Column (3) covers those who had been self-employed previously and those who are salaried on return.

The sub-samples differ in most respects. However, the most striking difference is that workers who take up self-employment after return have accumulated much larger savings (1086 dinars) on average when they were abroad than other workers of the restricted sample who are salaried after return (442 dinars) or than other workers from the full sample (466 dinars).

From the data we also find that those who took up self-employment mainly use their own capital for investment after return: 87.6% of projects are realized with savings accumulated during migration and only 12.4% of migrants received extra funds from special programs (Mesnard, 1999). None of them relied on formal bank credit, though some informal credit was probably available.[13] They explicitly mention their difficulties in getting access to credit markets when asked about the main obstacles faced in starting up their projects.

5. Estimation results

Since we are interested in whether a return migrant enters self-employment we restrict the sample to those who had not previously been self-employed prior to their migration. Table 2

[13] These findings echo those reported by Thomas (1992) using data for Lima, Peru.

gives the estimated parameters on the control variables. For the distributed lag we set m=10, although the results were quite similar with m=1.

The control variables are jointly significant, though only a couple of variables are individually significant.[14] Married respondents were less likely to start a new business; those born in the Center-East of Tunisia were more likely to do so. Individual enterprises have flourished in the Center-East region around Sousse, whose inhabitants are relatively mobile and have created networks with migrants working in France, Italy or Germany.

Figure 2(a) gives the nonparametric regression of $S_i - X_i \hat{\pi}_i$ on wealth, with its 95% confidence interval. We use the local regression method of Cleveland (1979).[15] The relationship is increasing and at least weakly concave over the whole range of the data. When we calculate the predicted probabilities we find that almost all (94%) of the sample is within the (0,1) interval; 6% of the predicted probabilities are negative, and none are above one. This is consistent with the assumed probability model, but is not a conclusive test (as noted in section 3). We test robustness to a (parametric) nonlinear probability model in the next section.

---

14 We find fewer significant non-wealth characteristics than in some past empirical work on the determinants of self-employment. We can only conjecture why. Possibly our sample is more homogeneous. Or possibly the significant controls in past work reflected the endogeneity of the wealth variable or miss-specification due to failure to allow for nonlinearity in the wealth effect. Possibly measurement error is complicating matters. (Wealth measurement errors will add an error component $\Delta[\phi(W_i^*) - \phi(W_i)]$ to (9), where W_i^* is true wealth, which would need to vanish in probability for consistency.) However, the accumulated savings of return migrants are probably known with precision.

15 This is often referred to as LOWESS (Locally Weighted Scatter Plot Smoothing) (Härdle, 1990, p.192). Deaton (1997) discusses the advantages of this method, and the closely related method proposed by Fan (1992), over kernel regression. We used two LOWESS programs as a cross-check (namely those in STATA and SAS; the latter gives confidence intervals directly). The results were very similar. We set the smoothing parameter to 0.8 with a linear local regressions; results were very similar for a 0.9 smoothing parameter and a quadratic local regression function.

Figure 2(b) gives the corresponding regressions without the controls. We also estimated these regressions without restricting the sample to those who had not been self-employed prior to migrating and we also tried dropping the extreme values for the richest individuals keeping the 80% poorest individuals in the sample; the results were very similar to Figure 2.

We find no evidence of increasing returns at low wealth. The curve is concave for low wealth levels, then becomes more linear for higher values. Adding control variables attenuates slightly the slope. These results suggest that it is possible in this setting to start some sort of business with little wealth, and that diminishing returns set in quickly. Thus higher wealth inequality reduces the rate of self-employment.

Our empirical results are consistent with a special case of our theoretical model in section 2. The special case assumes that labor cost is uniformly distributed, there are diminishing returns to capital throughout, and that capital is constrained by initial wealth. We cannot rule out other possible interpretations of our empirical results, though none appears more plausible than this special case of the model in section 2. For example, it is possible that wealth is picking up ability differences correlated with wealth, though our main findings are robust to adding controls for obvious ability correlates, such as education. Or possibly wealth is picking up differences in aversion to risk, though it would seem hard to explain how this could yield the concavity that we find. Amongst the variables we have in our data set, wealth appears to be of over-riding importance to whether or not a return migrant starts a new business, and the shape of the relationship is suggestive of the joint effect of borrowing constraints and diminishing marginal product of capital.

Our results are the opposite to those reported in Lindh and Ohlsson (1998), who concluded that lower wealth inequality in Sweden has reduced the rate of self-employment.

There are a number of differences in data and methods,[16] and it is difficult to speculate how these differences would affect the results. However, we can also remark that there is a seemingly plausible way of reconciling the two studies by noting that start-up capital requirements are likely to be considerably larger in Sweden than Tunisia.

6. Checks for robustness

Parametric nonlinear probability model

Could the lack of significant controls arise from violations of our assumption of a linear probability model in allowing for heterogeneity in the non-wealth characteristics? As a check on this possibility, we re-estimated the model as a probit with the wealth variable entering as a cubic function and the same linear controls as in Table 2 (except now the probability function is a normal distribution). Again the same two variables were significant at the 5% level as in Table 2, and no other variables were significant.

Selection bias

To test for selection bias in our results we pooled the return migrants with other non-migrant workers and estimated a probit in which the dependent variable took the value one if an observation was in the return migrant sample and zero otherwise. We then included the inverse Mills's ratio from this probit as one of the control variables in the vector X. The wealth variable entered linearly in the probit. While the model is identified on nonlinearity, we also used an instrument comprising a dummy variable for whether the worker was older than 18 in 1974. Our idea here is that reaching adulthood in 1974 creates a discontinuity in the probability of

---

[16] Notably in the fact that we have used micro data for a specific population group for whom wealth can reasonably be treated as exogenous, while Lindh and Ohlsson used aggregate time series data treating wealth distribution as exogenous.

migrating after that date, but does not affect the probability of starting a new business on return.[17] Otherwise the control variables in X were the same.[18] The probit revealed a number of significant factors in explaining who was a return migrant in the sample, and wealth, X and the instrument were jointly significant at better than the 0.00005 level.[19]

However, we found no evidence of selection bias. The Mill's ratio from the first stage probit was insignificant as an additional control variable; the t-test was 0.239 (0.395 if we drop the dummy for having reached adulthood in 1974 out of the selection equation). So this does not suggest that we need be concerned about selection bias in our estimates of the wealth effect on self-employment. (Recall that we cannot rule out selection bias with respect to migrants as a whole, since we do not have data on those that did not return to Tunisia; however, we are not aiming to draw inferences about migrants as a whole.)

Nonseparability

The model in section 2 follows the literature in assuming that start-up capital is solely a function of wealth (though we relaxed the assumption that it is a constant proportion of wealth). A more general specification allows worker characteristics to influence start-up capital independently of wealth. For example, it may be conjectured that better education allows a worker to borrow more at given wealth. Individual characteristics might enter either the production function or through differences in r. To allow this we consider the model:

---

17 Most Tunisian workers observed in the survey have migrated individually when reaching adulthood (Mesnard, 1999). Moreover, after 1974, host countries closed their borders to labor migration following political and economic changes including rising unemployment (Zimmermann, 1994).

18 One minor difference is that we used age in 1986 instead of age at the date of return, for consistency across the samples.

19 The significant variables at the 5% level were wealth (positive), the instrument (positive), age squared (negative), no education, secondary and higher education level (negative, relative to primary), number of dependents (positive), and being born in the center west or south east (positive).

$$S_i = \phi(W_i + X_i\gamma) + X_i\pi + \nu_i \tag{12}$$

To test this against (7) we take a first-order Taylor series expansion of the ϕ function:

$$\phi(W_i + X_i\gamma) = \phi(W_i) + \phi'(W_i)X_i\gamma + residual \tag{13}$$

This is of course an approximation. The residual in (13) includes higher-order effects that need not be innocuous. However, significant interaction effects between the controls and the estimates slopes of the ϕ function would at least be suggestive of misspecification.

Our test is thus to run the regression:

$$S_i - X_i\hat{\pi} - \phi(W_i) = \phi'(W_i)X_i\gamma + \nu_i \tag{14}$$

We could not reject the null hypothesis that $\gamma = 0$; indeed, the extra variables were only significant at the 67% level. So individual characteristics do not appear to interact with wealth in the liquidity constraint. We also estimated the model on the full sample (including those who had been self-employed prior to migrating); the results were very similar.

7. Implications

To interpret our results, we estimate the contribution of wealth inequality to the average rate of business starts amongst return migrants. This is given by:

$$\Delta \equiv \phi\left(\sum_{i=1}^{n} W_i / n\right) - \sum_{i=1}^{n} \phi(W_i) / n \tag{15}$$

which is positive for ϕ concave. Using the empirical non-parametric regression function, the value of Δ is 4% points. With complete equalization of wealth, the predicted rate of new business starts at mean values of the controls is 23.9%, as compared to a predicted mean on the same sample of 19.9%. This must be judged a modest impact given the extent of the wealth redistribution required.

As an aside, we repeated the above analysis using instead a probit for the probability of self employment and a cubic function of wealth (as mentioned in the last section). The three terms in the polynomial were all significant at the 5% level. On using this model to simulate the effect of equalizing wealth, the predicted mean probability of self-employment rose from 21.0% to 23.9%, very close to the results with the nonparametric model.[20]

A smaller redistribution naturally gives an even lower impact. Suppose that a redistributive wealth tax is introduced. The individual tax rate is an increasing function of initial wealth levels, and is budget neutral, i.e.,

$$t_i = T_i / W_i = \alpha + \beta W_i \tag{16}$$

where $\beta > 0$. When the value of α is set to assure that $\sum T_i = 0$ we have:

$$t_i = \beta(W_i - MSW / \overline{W}) \tag{17}$$

where $MSW = \sum W_i^2 / n$ is mean squared wealth and $\overline{W}$ is mean wealth. The value of β determines how much redistribution is achieved. To interpret β suppose that the pattern of taxes and subsidies given by (17) entails tax rate t^* on the richest person in the sample, with wealth W^*. Then:

$$\beta = \frac{t^*}{W^* - MSW / \overline{W}} \tag{18}$$

Using the estimated (individual-specific) first derivatives of S we can estimate the change in the probability of starting a new business for each individual under the hypothetical wealth

20 When we formed the predicted probabilities and plotted them against wealth the function traced out a sigmoid curve within the range of the data. However, this parametric from was sensitive to deleting extreme values. When we dropped the highest 1% of the sample in terms of wealth, the function became strictly concave over the range of the data. In this case the effect of redistribution was much smaller, with only a one percentage point gain in the average self-employment rate.

redistribution. Adding up these probability changes, we then have a measure of how sensitive the aggregate rate of self-employment is to the distribution of wealth. (This is not of course a policy simulation; rather it is a means of assessing how responsive the overall rate of business starts is to changes in inequality.)

Table 3 gives the percentage changes in the number of business starts for various wealth redistributions, indicated by the mean tax rate on the richest individual. As expected, the impact on aggregate self-employment rises with the extent of the redistribution. However, the effect of tax-subsidy scheme entailing as much as a 50% tax on the richest wealth-holder only increases the average rate of business starts by 0.61% points.

8. Conclusions

We have studied one of the arguments that has been made in the recent theoretical literature linking aggregate economic activity to the distribution of wealth. In our theoretical model, the distribution-dependence of the level of self-employment in the economy is potentially complex. On the one hand, diminishing returns to capital will tend to mean that greater wealth inequality yields a lower number of business start-ups at any given mean wealth in the economy. On the other hand, non-convexities in employment opportunities at low levels of wealth will tend to mean that inequality is good for aggregate business activity. The outcome is an empirical question.

In attempting to allow for this potentially complex distribution dependence, we have not imposed a specific functional form on the data, but have drawn instead on flexible nonparametric methods to study the empirical relationship between initial wealth and new business start ups. In the hope of better identifying the relationship, we have focused solely on return migrants in a developing country who have brought back diverse amounts of accumulated savings from their period abroad, and can be expected to be contemplating whether to take up self-employment on returning to their home country.

Our data on return migrants in Tunisia are consistent with the joint effect of borrowing constraints and diminishing returns to capital. We do not find any sign of increasing returns at low wealth — contrary to what one would expect if there were sizable start-up costs. The wealth effect is positive and at least weakly concave over the range of the data. Our results imply that, at any given mean, the higher the initial inequality of wealth, the lower the overall rate of own-business start-ups.

However, the quantitative magnitude of the inequality effect is modest. Even full equalization of wealth at a given mean would only increase the rate of new business starts amongst return migrants from 20% to 24%. Less redistribution will naturally have less impact. A progressive redistribution entailing a tax rate that rises with wealth, reaching 50% for the richest person, would increase the rate of self-employment by less than one percentage point.

These calculations are at best suggestive. They are not policy simulations, since they do not take account of any incentive costs of redistribution. However, one cannot presume that such costs would be positive; for similar reasons to why lack of wealth constrains the ability to start a new enterprise, it may well reduce its productivity. Possibly redistribution would have larger effects on aggregate profitability than we have found for business start-ups.

References

Aghion, Philippe and Patrick Bolton, 1997, "A Trickle-Down Theory of Growth and Development with Debt Overhang," *Review of Economic Studies* 64(2): 151-62.

Banerjee, Abhijit V., and Andrew F. Newman, 1993, "Occupational Choice and the Process of Development," *Journal of Political Economy* 101(2): 274-298.

Bates, Timothy, 1990, "Entrepreneur Human Capital Inputs and Small Business Longevity", *Review of Economics and Statistics*, 72(4): 551-559.

Bhattacharya, Utpal and B. Ravikumar, 2001, "Capital Markets and the Evolution of Family Businesses," *Journal of Business* 74(2): 187-219.

Blanchflower David G., and Andrew J. Oswald, 1998, "What Makes an Entrepreneur? Evidence on Inheritance and Capital Constraints," *Journal of Labor Economics* 16(1): 26-60.

Cleveland, William S., 1979, "Robust Locally Weighted Regression and Smoothing Scatter Plots," *Journal of the American Statistical Association* 74: 829-36.

Deaton, Angus, 1997, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: Johns Hopkins University Press for the World Bank.

Dunn, Thomas and Douglas Holtz-Eakin, 2000, "Financial Capital, Human Capital and the Transition to Self-Employment: Evidence from Intergenerational Links," *Journal of Labor Economics* 18(2): 282-305.

Evans, David S. and Boyan Jovanovic, 1989, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy* 97(4): 808-827.

Evans, David S and Linda S. Leighton, 1989, "Some Empirical Aspects of Entrepreneurship", *American Economic Review* 79 (3): 519-535.

Fan, Jianqing, 1992, "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association* 87: 998-1004.

Hall, Peter, J.W. Kay, and D.M. Titterington, 1990, "Asymptotically Optimal Difference-Based Estimation of Variance in Nonparametric Regression," *Biometrika* 77: 521-28.

Härdle, W., 1990, *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.

Holtz-Eakin, Douglas, David Joulfaian. and Harvey S. Rosen, 1994, "Entrepreneurial decisions and liquidity constraints," *RAND Journal of Economics*, 25(2): 334-347.

Ilahi, Nadeem, 1999, "Return Migration and Occupational Change," *Review of Development Economics*, 3(2): 170-186.

Kanbur, Ravi, 1979, “Impatience, Information and Risk Taking in a General Equilibrium Model of Occupational Choice,” *Review of Economic Studies* Vol.46(4): 707-18

Kihlstrom, Richard E., and Jean-Jacques Laffont, 1979, “A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion,” *Journal of Political Economy* 87 (August): 719-748.

Levy, Brain, 1993, “Obstacles to Developing Indigenous Small and Medium Enterprises: An Empirical Assessment,” *World Bank Economic Review* 7(1): 65-83.

Lindh, Thomas and Henry Ohlsson, 1996, “Self-Employment and Windfall Gains: Evidence from the Swedish Lottery,” *Economic Journal* 106: 1515-26

Lindh, Thomas and Henry Ohlsson, 1998, “Self-Employment and Wealth Inequality,” *Review of Income and Wealth*, 44(1), 25-42.

Magnac, Thierry and Jean-Marc Robin, 1996, "Occupational choice and liquidity constraints" *Ricerche Economiche* 50, 105-133.

Matzkin, Rosa L., 1992, “Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models,” *Econometrica* 60(2): 239-270.

Mesnard Alice, 1999, "Migration Internationale, Accumulation D'épargne et Retour des Travailleurs," PhD. dissertation, Ecole des Hautes Etudes en Sciences Sociales, Paris.

Paulson, Anna and Robert Townsend, 2000, "Entrepreneurship and Liquidity Constraints in Rural Thailand", mimeo, University of Chicago.

Thomas, James J., 1992, *Informal Economic Activity*, New York: Harvester Wheatsheaf.

United Nations, 2000, *The World’s Women 2000: Trends and Statistics*, New York: Department of Economic and Social Affairs.

Van Praag, C.M., and H. Van Ophem, 1995, “Determinants of Willingness and Opportunity to Start as an Entrepreneur,” *Kyklos* 48: 513-40.

Yatchew, Adonis, 1997, “An Elementary Estimator of the Partial Linear Model,” *Economics Letters* 57: 135-143.

Yatchew, Adonis, 1998, “Nonparametric Regression Techniques in Economics,” *Journal of Economic Literature* 36(June): 669-721.

Zimmermann, K.F., 1994, "Immigration Policies in Europe : An Overview" in Siebert, H (ed.), *Migration: A Challenge for Europe*, J.C.B. Mohr: Tubingen.

Figure 1: Theoretical relationship between business start-ups and wealth

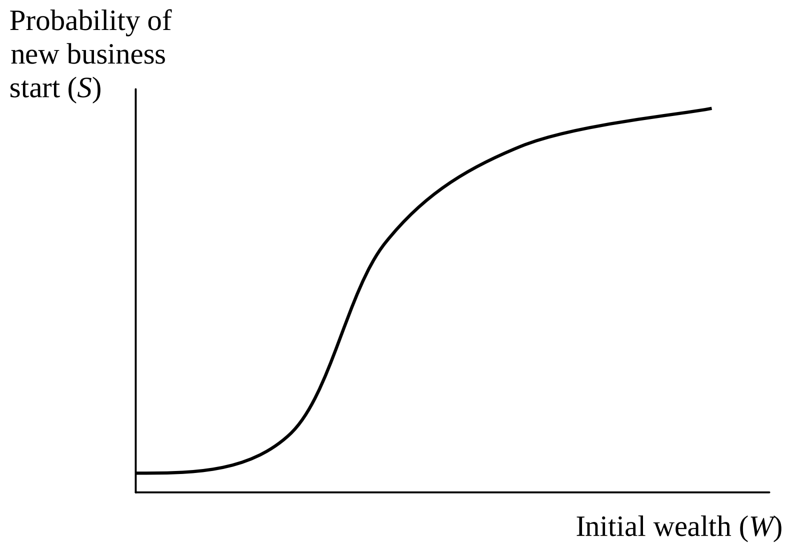
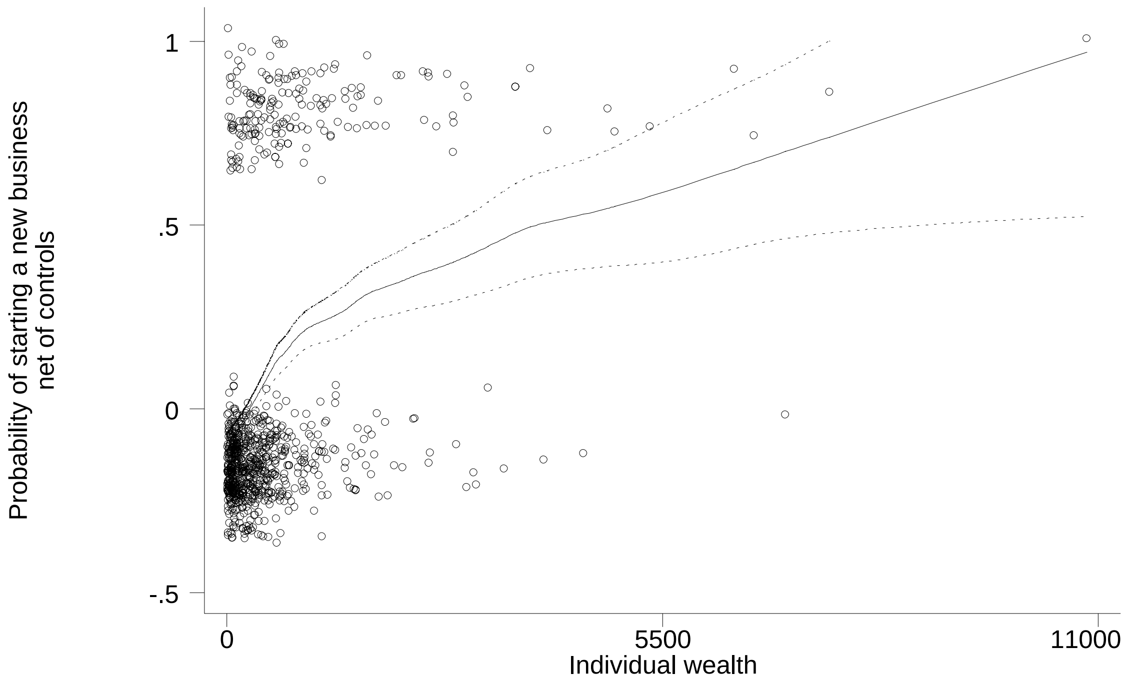


Figure 2: LOWESS regression of business starts on initial wealth with and without controls (including 95% confidence intervals)

(a) With controls; dependent variable is $S_i - X_i \hat{\pi}$.



(b) Without controls

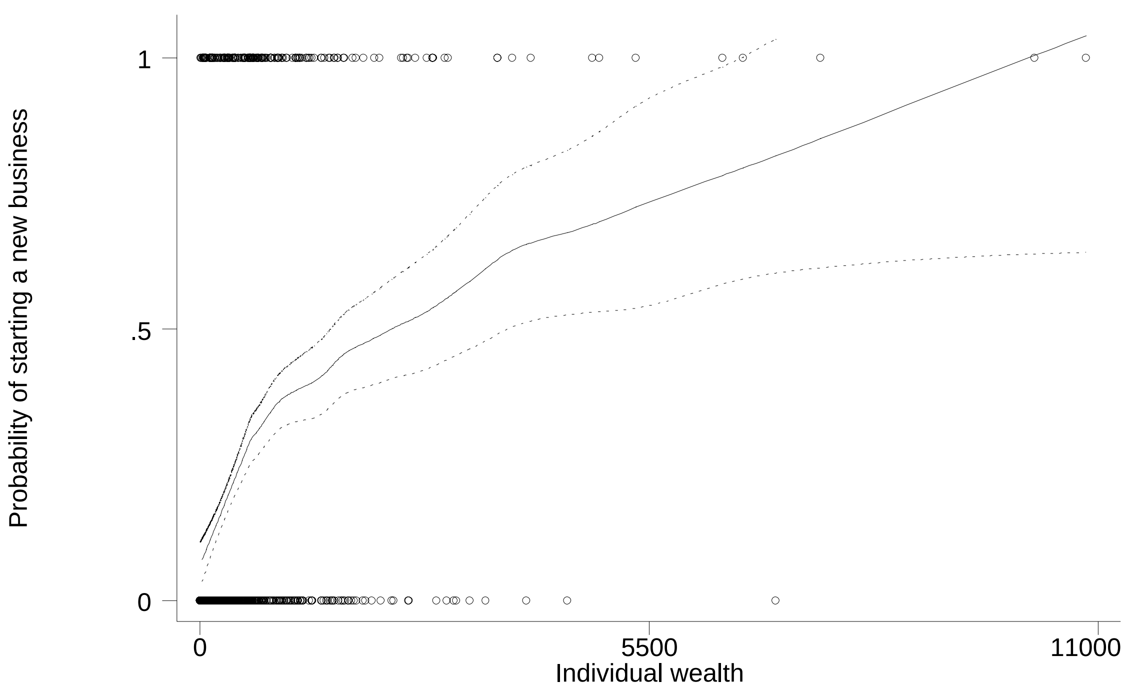


Table 1 : Descriptive statistics

	Not previously self-employed			
	(1) Took up self-employment (n=210)	(2) Others (n=840)	(3) Rest of full sample (n=1014)	(4) Total sample (n=1224)
Age at return	34.87	32.44	33.66	33.87
	(10.60)	(10.66)	(0.38)	(11.13)
No education (%)	36.4	33.3	37.7	37.5
Primary school (%)	46.4	50.2	46.4	46.4
Short secondary school (%)	4.3	4.7	4.5	4.4
Long secondary school (%)	12.9	11.8	11.4	11.7
Number of dependents	5.0	4.55	4.79	4.82
	(2.94)	(2.93)	(2.96)	(2.96)
Married (%)	81.4	80	81.6	81.6
Born in area of Tunis (%)	5.3	4.9	4.4	4.6
Born in Center East (%)	23.3	18.6	20.5	21
Born in Center West (%)	21	24	24.4	23.8
Born in Southern East (%)	15.7	23.1*	22.3*	21.1
Born in Southern West (%)	10	10.6	10.2	10.1
Born in Northern East (%)	7.6	6.1	5.6	6
Born in Northern West (%)	17.1	12.7	12.6	13.4
Savings at return	1086.2	442.35*	465.9*	580.10
	(1539.13)	(951.77)	(926.45)	(1091.94)

Note: * significantly different from the mean in column (1), t-test.

Table 2 : Parameters on control variables in explaining the probability of starting a business amongst return migrants

	Coefficient (t-ratio)
age at return	0.011
	(1.84)
age squared	0.000
	(1.5)
no education	0.053
	(1.2)
short secondary school	-0.103
	(1.17)
long secondary school	-0.021
	(0.43)
Married	-0.133
	(2.96)**
number of dependents	0.003
	(0.40)
born in Center East	0.107
	(1.88)*
born in Center West	-0.009
	(-0.18)
born in Northern East	-0.003
	(0.05)
born in Northern West	0.013
	(0.22)
born in South East	-0.067
	(1.28)
born in area of Tunis	-0.07
	(0.94)
Constant	-0.002
	(0.13)
Observations	695
R-squared	0.04
Tests of $\gamma = 0$: F(13,869)	0.79
	Prob=0.67

Note: Robust t-statistics in parentheses; one extreme value (savings=20550) dropped.
*significant at 5% level; ** significant at 1% level.

Table 3: Effects of wealth redistribution on the number of business start-ups

t^*	Absolute change (% points)	% change ($\bar{S}$ =19.85%)
5%	0.06	0.30
10%	0.12	0.60
25%	0.31	1.56
50%	0.61	3.07

Note: The table gives the predicted changes in the proportion of previously not self-employed return migrants who start a new business as a result of a progressive tax/subsidy scheme such that the richest individual has a tax rate of t^*.