The Wealth Effect on New Business Startups in a Developing Economy

Alice Mesnard — *Institute of Fiscal Studies, London*

Martin Ravallion[1] — *World Bank, Washington DC*

September 2004; revised April 2005

Abstract: Various theoretical models have postulated nonlinearities in the wealth effect on self-employment stemming from start-up costs and/or liquidity constraints. Nonlinearity implies that the extent of entrepreneurial activity in an economy depends on the distribution of wealth, though in potentially complex ways. To test for nonlinearities, we estimate both a non-parametric linear probability model and a parametric nonlinear model of the choice to be self-employed amongst return migrants in Tunisia. Controls for heterogeneity are included, and tests are made for selection bias and separability between wealth and the controls. The relationship between the probability of starting up a business and savings repatriated at return is concave for almost all the range of our data, though we find weak evidence of a convex relationship at very low wealth levels. Our results indicate that the aggregate self-employment rate is an increasing function of aggregate wealth, but a decreasing function of wealth inequality.

JEL: D31, M13

Keywords: Borrowing constraints, wealth inequality, self-employment, migration

---

[1] This is a revised version of an earlier working paper, "Wealth Distribution and Self-Employment in a Developing Country," Center for Economic Policy Research DP 3026, October 2001. 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 Erich Battistin, Bruno Biais, Ian Crawford, Pierre Dubois, Esther Duflo, Emanuela Galasso, Robert Townsend, Vassilis Hajivassiliou, Peter Lanjouw, Michael Lokshin, Mans Soderbom, Dominique van de Walle, Frank Windmeijer and seminar participants at the Institute of Fiscal Studies, London, and the Center for the Study of African Economies, Oxford. We also thank a referee for valuable suggestions.

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 <u>level</u> of aggregate wealth matters to aggregate self-employment. But does the <u>distribution</u> 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.[4]

---

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).

4 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 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 tests for nonlinearity in wealth effects on business start-ups 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.[5] 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.[6] 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 an endogenous credit limit, given limited commitment.

5 One can conjecture that credit constraints also impede business development and longevity, and there is some supportive evidence (Bates, 1990; Levy, 1993; Woodruff and Zenteno, 2001). However, we do not know of any evidence that wealth inequality matters to aggregate profitability and/or longevity.

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

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 in developing countries than developed ones, our theoretical model suggests that wealth inequality can have opposite effects on employment structure between the two types of countries.

Motivated by this model, we look for non-linear 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. 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). They found a positive correlation between self-employment and inequality, which was robust to de-trending the data.

We use instead micro data on self-employment and individual wealth and we control for other individual, family and regional characteristics, in line with the recent literature on entrepreneurship and liquidity constraints. A concern about past work is that wealth data collected at the date of survey may be endogenous to occupational choices. A few empirical studies overcome this problem by studying the transition from wage employment to self-employment (rather than self-employment per se) and combining this with data on wealth and other characteristics at a time prior to this transition (Evans and Jovanovic, 1989, Paulson and Townsend, 2000), which are considered exogenous; examples include inheritance or gifts (Blanchflower and Oswald,1998, Holtz-Eakin, Jouflaian and Rosen,1994) or windfall gains (Lindh and Ohlsson, 1996). Studying return migrants and their wealth at the date of return serves a similar function.[7] There is a natural dynamic sequence between the prior wealth accumulation of migrants and their occupational choices on return. The fact of being predetermined does not of course guarantee exogeneity; possibly some latent propensity for self-employment prompted greater wealth accumulation while abroad. 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. We are also able to test exogeneity using for identification a feature of the history of European migration policy.

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. A further difference with past work is that we treat the wealth effect in a flexible way using a partial linear model, although we test robustness using a parametric non-linear probability model.[8] Our results show that wealth increases the probability of starting up a business, but with diminishing returns over most of the data. We also offer a test for whether other characteristics interact with wealth in the liquidity constraint.

The following section discusses the partial equilibrium effect of wealth on business start-ups in general terms, while section 3 outlines our econometric methods. Section 4 describes the

---

7 We know of a few studies on occupational choices of return migrants in developing countries, which found that higher accumulated savings while abroad are associated with higher probabilities of entering self-employment relative to wage labor (Mesnard, 1999, 2004; Ilahi, 1999; Dustmann and Kirchkamp, 2002).

8 A few studies allow for non-linearities in the wealth effect on occupational choice using parametric estimation methods. For example, Evans and Jovanovic (1989) enter past wealth in a quadratic form and find a jointly significant effect of the linear and quadratic terms but wealth squared is not individually significant at conventional levels. Paulson and Townsend (2000) find that the coefficient on the wealth squared is significant and negative, but very small, for the whole sample and not significant for the Northeast and Central region of Thailand.

data and gives some descriptive results. Section 5 presents a test for possible endogeneity problems that may cloud the wealth self-employment relationship and shows that exogeneity of savings accumulated abroad cannot be rejected. Under such an exogeneity assumption, Section 6 presents the estimation results for a nonparametric linear probability model of self-employment and, in order to assess their robustness, provides results for an alternative parametric nonlinear probability specification . Section 7 gives our conclusions.

2. Nonlinear wealth effects on self-employment in theory

We take the distribution of wealth as given.[9] The amount of accumulated savings is W_i and is bounded below by $W_{\min} (\geq 0)$ and above by $W_{\max}$. With these savings, person *i* can 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.,

$$K_i / W_i = k \geq 1 \text{ for all } i \qquad (1)$$

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.

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 $K^* > 0$). 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 leave the form of the production function flexible.

---

[9] Mesnard (2001) provides a more general theoretical model of migration with endogenous wealth distribution.

There is a fixed own-labor requirement 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). We treat ω_i as a random variable with continuous twice-differentiable distribution function Ψ. (Note that ω_i also includes idiosyncratic 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:

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

with slope:

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

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 will depend on 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 mean-preserving 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*:

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

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. (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. Estimation methods

To investigate empirically the wealth effect on self-employment we use both non-parametric and parametric regression methods. In anticipation of our empirical results of section 5 we shall treat wealth as exogenous to occupational choice.

The first method we use is a non-parametric regression that requires very little structure or statistical assumptions in estimating the wealth effect. We also allow for heterogeneity in non-wealth characteristics as 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 we assume a linear probability model with linear controls. Our method thus entails estimating partial linear regressions, in which the sub-function for the wealth effect is kept completely flexible, through non parametric estimation.

We write the probability of becoming self-employed as some unknown function of *W* and a linear function of the 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. (We compute robust standard errors to allow for possible heteroskedasticity.) 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.

To estimate equation (7) we draw on the literature on partial linear models following Robinson (1988) (as reviewed by Yatchew, 1998). We order all observations by 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 ($\text{plim}[\phi(W_i)-\phi(W_{i-1})]=0$)[10]. 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.

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:

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

10 We checked that the actual values of *W* are sufficiently close to each other, so that equation (9) is approximately true for "adjacent" observations.

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, though under the null hypothesis none of these effects matter. Indeed, under the null hypothesis, all the parameters can be consistently estimated, as if the true model was (7) and the estimation proceeds along the line described above (we first estimate π, then ϕ, and we interact X_i with ϕ' as shown in (13)). Therefore significant interaction effects between the controls and the estimates slopes of the ϕ function would at least be suggestive of misspecification. Our test for separability between *W* and *X* is thus to run the regression[11]:

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

Linearity in the control variables has the well-known shortcoming that there is no guarantee that the predicted values for *S* will be in the [0,1] interval. We will check if our estimates satisfy this condition. However, even if satisfied at all data points, that check is not conclusive since the parameter estimates on which it is based will be inconsistent if the underlying probability model is miss-specified. 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.[12] To check robustness, we employ a second method using a parametric nonlinear probability function.

---

11 However it is worth noting that there are concerns about the statistical power of this procedure, as, for example, the residual in equation (13) and therefore in equation (14) would depend on X_i under the alternative.

12 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.

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.[13] The survey was conducted in all geographical areas of Tunisia (both rural and urban areas) covering return migrants and the other non migrants Tunisian workers.

Return migrants are defined as workers who have worked abroad at least once during 1974-86 and returned to live in Tunisia before the survey date. This group was deliberately over-sampled. 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.

The non-migrant sample suggests a correlation between wealth and self-employment. From the survey we can measure the (monetary and in kind) wealth accumulated from 1974 until 1986. Similarly to the return migrant sample, we dropped from the non-migrant sample the workers who were self-employed before 1974. Table 1 shows that the mean wealth of those who became self-employed between 1974 and 1986 is 689 dinars, while it is 397 for those who did not (to be compared to the GDP per capita which was about 980 dinars in 1986). The difference is statistically significant (with a t-test of 2.02).

It could be hard to detect credit-constraints in the non-migrant sample. We know whether the respondent was self-employed at the date of interview. And we know wealth accumulation up to that date. However, there are clearly serious concerns about the endogeneity of wealth with respect 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).

---

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

We focus instead on the sample of return migrants. Returning from a long period overseas makes a natural break in work history. Such migrants make a choice as to whether or not they will start a new business on returning to the home country. And we can identify how much they brought back from their period overseas — hereafter called “savings” — 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.

Nonetheless, there are various potential sources of bias to consider. Migrants may not be typical of the workforce in the origin country and return migrants can be thought of as a sample selected from a complete set of migrants, not all of whom returned. In particular, temporary migrants may be selected on their wealth level and abilities to accumulate wealth abroad, if migration is a way to overcome liquidity constraints in the origin country, as argued by Mesnard (2004). It is plausible that the probability of returning depends positively on accumulated wealth. The sample selection process is unknown and we have no data on those who did not return. This would also explain why return migrants have accumulated more savings on average than non migrants, as shown in Table 1.[14] Then the wealth effects we see in the data on return migrants also include an unknown selection effect, as well as liquidity constraints. We will address this issue in the next section that proposes a test for exogeneity of savings

Our main sample covers 1050 male returned migrants who reported that they intend to stay indefinitely in Tunisia.[15] The survey obtained general information about their migration

[14] Note that the variable measuring “savings” is not the same across samples since migrants returned from migration before the survey. Depending on the life-cycle pattern of saving accumulation, the difference between savings accumulated abroad measured for return migrants and savings accumulated at the date of survey for non migrants may also underestimate or overestimate the difference between savings that would be measured at the same date of survey for the two samples.

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

history (number of migrations, dates, locations, return motives, duration, employment) and their working and living conditions during their last migration. To identify new business start ups 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.[16]

The survey obtained data on a number of obvious control variables, including age and education. Whether one takes up self-employment may also be correlated with where one lives, as an influence on proximity to markets. However, endogeneity concerns speak against controlling for current location in this context (given that Tunisian workers can in principle choose where they return to). Also self-employment rates and amounts of savings brought back to Tunisia vary significantly with the country of migration. In particular, migrants returning from France or other European countries tend to have accumulated more savings and are more often self-employed than migrants returning from Libya or other Arabic countries (for further details see Mesnard, 1999). Once again, it is difficult to consider the country of migration as being exogenous. We do include controls for place of birth.

The data give some information on the family structure, which is also likely to play a role in the decision to start-up a business after return (through better access to informal sources of credit, by providing the return migrants with cheap labour force, or by offering job opportunities in family-type enterprises). We use family size to capture these effects.

[16] There is also information on pre-return transfers sent by migrants via remittances, which are described in detail in Mesnard (1999). However we chose not to use this proxy of wealth since data on transfers suffer from obvious miss-reporting and there are many non random missing answers (only 83 return migrants give an answer). Also it is not clear which part of these transfers could be used as capital to start up a business after return.

Table 1 provides summary statistics. Column (1) gives data on the sample of workers who had not been self-employed prior to migrating but took up self-employment on return. Column (2) gives data on those who continued to be salaried on return. The full sample of return migrants is described in Column (3).

The sub-samples differ in most respects. 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 who are salaried after return (442 dinars). This is clear when comparing the distributions of savings of the two sub-samples, as represented by Figure 2.

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.[17] They explicitly mentioned their difficulties in getting access to credit markets when asked about the main obstacles faced in starting up their projects.

5. Test for the exogeneity of savings

As a precursor to the main analysis, we first address the aforementioned concern that savings might be endogenous. This could arise from unobserved heterogeneity or more structural reasons linked to the simultaneous decisions made by temporary migrants. To test this, we replicated the two step Probit estimation method with instrumental variables *a la* Rivers and Vuong (1988) that is presented in detail in Mesnard (2004).

---

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

The identification strategy proposed in Mesnard (2004) rests on a distinction between migration that occurred before 1974 and that after. This discontinuity in migration history determines the different levels of savings brought back to Tunisia by migrants. Migrants did not have the same migration opportunities if they migrated before 1974 than after, given very different political and economic situations. In particular, European countries closed their borders after 1974, which led many Tunisian workers to migrate towards Libya where wages were lower than in European countries. And those who still migrated to European countries after 1974 were encountering more difficult conditions in the destination labour markets.

We require two assumptions in using this aspect of migration history for identification purposes. First, that there are no important cohort differences in latent entrepreneurial activity beyond what the age control picks up; second, that this turning point in migration history at 1974 was not anticipated by Tunisian workers and that the migration date is exogenous. This is consistent with our data showing that Tunisian workers of our sample tend to migrate at the beginning of their working life.[18] The results of the instrumented regressions presented in Table 2 confirm that the chosen identifying instrument determines significantly and strongly the savings accumulated abroad by temporary migrants, as can be seen from the coefficient for "having migrated before 74."

Based on the above, we found that exogeneity of savings is not rejected by our test: the residual of the instrumented regressor for savings is not significant at conventional levels when added to the control variables in the main regression as shown in Table 3.

[18] Note that we could not use either the country of migration itself, migration duration, wages in foreign country or activity abroad as identifying instruments, and could not find any additional instruments, which would have allowed us to test for over-identifying restrictions. All such variables based on the migration history of migrants are most likely to be endogenous.

On this basis we decided to treat migrants' savings accumulated abroad as exogeneous to their decision to take up self-employment or not on their return.

6. Results for the nonparametric linear probability model

We can now study the non-linearities in the wealth effect on self-employment using the partial linear non parametric method presented in section 3 under the exogeneity assumption discussed in the previous section. Table 4 gives the estimated parameters on the control variables.[19]

The control variables are jointly significant, though only a few variables are individually significant. We find that 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.

In testing separability between wealth and the controls, we could not reject the null hypothesis that $\gamma = 0$ (equation 14); indeed, for return migrants, the extra variables were only significant at the 73% level. In sum, individual characteristics do not appear to interact with wealth in the liquidity constraint.

Figure 3(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).[20] The relationship

---

[19] We dropped one outlier (the richest individual) from the sample of return migrants. Furthermore, although in the main analysis we set m=10 for the distributed lag, we also performed a sensitivity analysis with respect to different choices of this parameter finding no main changes.

[20] 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

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 will also test robustness in relation to a (parametric) nonlinear probability model below.

We also estimated non parametrically the bivariate relationship without controls, using straightforward LOWESS regression methods, as shown in Figure 3(b). We find the same concave relationship over the whole range of data. Adding control variables attenuates slightly the slope.

For comparison, we also performed the same estimations using the sample of non migrants, for which savings are not measured before the occupation is chosen. The results were very similar.[21] Again there was a concave effect of accumulated savings on the probability to start up a business. The other controls do not appear to interact with wealth in the liquidity constraint (in the separability test, the extra variables were only significant at the 69% level), as shown in Table 4.

We also estimated these regressions without restricting the sample to those who had not been self-employed prior to migrating and we tried dropping the extreme values for the richest individuals, keeping the 99% poorest individuals in the sample. In addition we experimented with different bandwidths. The results remained very similar.

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.

[21] The results on the control variables for the sample of non migrants are presented in Table 4. We do not present the graphs of the wealth self-employment relationship since they are very similar to Figures 3(a) and 3(b).

In order to get a more accurate idea of the relationship for the lower bound of the wealth distribution where there is a concentration of our sample, we re-estimated the relationship keeping the 95% poorest individuals (hence dropping individuals with savings higher than 2000 dinars). Results with the control variables are presented in Figure 4(a) and in Figure 4(b) without controls. We now find some weak evidence of a convex relationship at very low wealth levels, although we cannot reject the concavity of the relationship at the 5% level. Subsequently the curve becomes concave for almost all the range of the data. These results suggest the existence of liquidity constraints but that diminishing returns set in quickly.

We find fewer significant non-wealth characteristics than in some past empirical work on the determinants of self-employment.[22] We conjecture that this is either because our sample is more homogeneous or because 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.

One difference with past work is our use of the nonparametric partial linear model. As noted in Section 3, this has the limitation that it restricts us to a linear probability model. To test robustness, we also estimated probits with the wealth variable entering as a polynomial function up to degree six and the same linear controls as Table 4. (Note that the functional form outlined in Figure 4 can be suitably approximated by a polynomial of some order.) We found that wealth entered as a cubic function and the same variables were significant at the 5% level, as reported in column (1) of Table 5. However, this is not a robust feature of the parametric model. In particular, we found that the cubic term in wealth was sensitive to deleting extreme values. When

[22] We also tested for the joint significance of age and age squared, as well as for the education variables, which were rejected at high test levels.

we dropped the highest 1% or highest 5% of the sample in terms of savings, the function became strictly concave over the range of the data and the same variables were significant at conventional levels, as shown in column (3) in Table 5 and column (1) in Table 6. This reversal does not occur in the nonparametric regression. For the non-migrant sample, the parametric method gives very similar results to Table 4, as shown in column (2).

In the same parametric specification, we also checked for the specification of the liquidity constraint by testing the joint significance of the interacted terms of the control variables with the wealth variable. The interaction effects were jointly insignificant at the 5% level. This helps address our worry that the nonparametric test of the specification for the liquidity constraint may suffer from power problems.

It is also possible that wealth is picking up ability differences correlated with wealth, which would cloud the relationship between startup wealth and self-employment. Although we have controlled for some obvious ability correlates, such as education, these measures do not distinguish entrepreneurial ability from ability in the labor market. As a robustness check, we added several additional control variables that are likely to be correlated with the unobserved entrepreneurial abilities of workers, namely whether the migrant acquired a qualification abroad, whether he had a skill before migrating and whether he was unemployed before migrating. We recognize that these additional controls are likely to be endogenous to occupational choice at return. However, they do at least allow us to offer a partial test of the robustness of our main results to omitted ability attributes. The results are in Table 6. As shown in column (2), the effects outlined above are remarkably robust to adding the extra control variables. We find that having acquired a skill abroad increases significantly the probability of becoming self-employed after returning. We also tried adding the duration of unemployment upon return (see in column

3), which might also pick some heterogeneity in abilities of return migrants. We found a significant negative correlation between unemployment duration at return and self-employment; again there are concerns about possible endogeneity. Nonetheless, our key results are quite robust to these tests.

A further concern is that since all return migrants did not return in 1989, the wealth effect might be picking up a time effect, in that entry into self-employment could depend also on how long the migrant had been in Tunisia. We added to the control variables the duration of time spent in Tunisia since return, but found no significant effect, as shown in column (4). Other results are very similar as in Table 5.[23]

It might also be argued that workers returning from different countries have different abilities, to the extent that the difficulties in migrating to European countries are greater than for Libya (the main destination country chosen by 77% of the sample) and workers may choose their migration destination depending on unobservable characteristics. Another justification is that migration costs are likely to be different and workers choosing to migrate to European countries may have a greater family wealth, which may cloud the relationship between startup wealth and self-employment.

To try to capture these unobservable sources of heterogeneity we tested robustness to including the migration country among the explanatory variables, trying several specifications as shown in Table 7. Again we have to keep in mind that there are endogeneity concerns about these extra controls (which is why we did not include these variables in our main results). We also estimated the relationship separately for the migrants returning from European countries versus Libya. Unfortunately we do not have enough observations to convincingly estimate the

[23] Note that the age at return becomes significant when adding the duration since return, which could simply pick up the effect of the duration spent abroad.

model with the 174 workers returning from European countries. Hence we only present the results in columns (3) and (4) in Table 6 for workers returning from Libya.[24] As we can see in columns (2) and (3) the effect of accumulated savings is very similar to our main results reported above and the same control variables are significant. On dropping the 12 richest individuals (by keeping the workers with savings less than 2000 dinars) we found once again a very significant quadratic wealth effect, as can be seen in column (4).[25]

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, that there are diminishing returns to capital, 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 differences in aversion to risk, although none of the return migrants surveyed in Tunisia mentioned related problems when they report the main obstacles faced in starting up their projects and it would seem hard to explain how this could yield the non linearity that we find. Amongst the variables we have in our data set, wealth appears to be of over-riding importance as 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 suggest further that higher wealth inequality, at least amongst individuals above a critical level of wealth, reduces the rate of self-employment**.** To interpret our results, we

---

[24] Note that the sample becomes too small to use semi parametric approaches.

[25] Keeping these extreme values in our sample the quadratic term is still significant, but with confidence level lower than 8% as shown in column (3).

can measure the contribution of wealth inequality to the average rate of business startups amongst return migrants in Tunisia. This is given by:

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

which is positive (negative) for ϕ concave (convex) and *a priori* ambiguous if ϕ is S shaped. Using the empirical non-parametric regression function, the value of Δ is 6.2% points. With complete equalization of wealth, the predicted rate of new business startups at mean values of the controls is 27.2%, as compared to a predicted mean on the same sample of 21%. This must be judged a modest impact given the extent of wealth equalization required

It should be emphasized that this exercise is not a policy simulation of the effect of wealth redistribution on the rate of business startups. We prefer to interpret Δ as a measure of the extent of the concavity in the empirical wealth effect rather than a policy simulation, which would need to also take account of any incentive effects. 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 affect its productivity (McKenzie and Woodruff, 2004). General equilibrium effects would also be relevant to a full assessment of the policy implications. A complete assessment of the effect of wealth redistribution on the aggregate rate of new business activity would require embedding the partial equilibrium relationship we have studied here into a dynamic general equilibrium model, which is beyond the scope of this paper. In such a model, wages in the labor market would be determined simultaneously with the occupational structure of the economy. Possibly, large flows of return migrants investing their savings in small enterprises in their country of origin would have long-run effects on wealth distribution and occupational structure in an economy with imperfect credit markets; further analysis of these issues can be found in Mesnard (2001).

Our results stand in contrast 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.[26] 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.

7. Conclusions

We have tested for nonlinearity in the wealth effect on new business start-ups in a developing country. The existence of nonlinearity is often assumed in theoretical work linking aggregate economic activity to the distribution of wealth. In our theoretical model of the relationship between entrepreneurship and wealth, 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 answer that question, we have focused mainly on return migrants in Tunisia 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. We have argued that

[26] Most notably, we have used micro data for a specific population group and, hence, controlled for a bunch of individual, family and geographic determinants while Lindh and Ohlsson used aggregate time series data and do not control for other possible determinants of the long-run evolution of employment structure of the economy.

accumulated savings while abroad can be treated as exogenous to the probability of starting a new business on return, and have provided an exogeneity test to support this assumption, exploiting an aspect of European migration history for identification.

Our results are consistent with the joint effect of borrowing constraints and diminishing returns to capital. The wealth effect is positive and at least weakly concave over the upper range of the data. Non parametric methods show weak evidence of non-convexity at very low wealth levels that could suggest the existence of a low but positive start-up costs for the projects realized by return migrants in Tunisia, although the concavity of the relationship cannot be rejected at a conventional confidence level. This is confirmed while using different parametric specifications that outline a quadratic wealth effect once we drop the richest individuals of our sample.

The estimated concave relationship we find suggests that wealth inequality reduces the aggregate level of business start-ups. Using our estimates to simulate the contribution of wealth inequality, our results imply that the higher the initial inequality, the lower the overall rate of own-business start-ups. The quantitative magnitude of the inequality effect seems small. Even in the extreme case of full equalization of wealth at a given mean, the rate of new business startups amongst return migrants would rise from 21% to only 27%. These calculations are at best suggestive. They are not policy simulations, since they do not take account of any incentive costs of redistribution or the general equilibrium effects, which would probably reduce the impact of feasible wealth redistribution on the rate of new business activity.

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.

Blundell R. and JL Powell (2003), Endogeneity in Nonparametric and Semiparametric Regression Models, in Dewatripont, M., LP Hansen, and SJ Turnovsky, eds , Advances in Economics and Econometrics, Theory and Applications, 8th World Congress, Vol 2 (Cambridge University Press).

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.

Dustmann, Christian and Kirchkamp, Oliver (2002), The optimal migration duration and activity choice after re-migration, *Journal of Development Economics* 67: 351-372.

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* 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, Brian (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.

McKenzie, David and Christopher Woodruff (2004), Do entry costs provide an empirical basis for poverty traps? Evidence from Mexican microenterprises, mimeo.

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.

____________ (2001), Temporary migration and intergenerational mobility,, *Louvain Economic Review,* 67(1): 59-88.

____________ (2004), Temporary migration and capital market imperfections, *Oxford Economic Papers*, 56 : pp.242-262.

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

Rivers D., and Vuong Q.H. (1988), Limited information estimators and exogeneity tests for simultaneous probit models, *Journal of Econometrics*, 39: 347-386.

Robinson P.M. (1988), Root-n consistent semiparametric regression, *Econometrica,* 56, 931-54.

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.

Woodruff, Christopher and Rene Zenteno (2001), Remittances and Microenterprises in Mexico, UCSD, Graduate School of International Relations and Pacific Studies Working Paper.

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

______________ (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

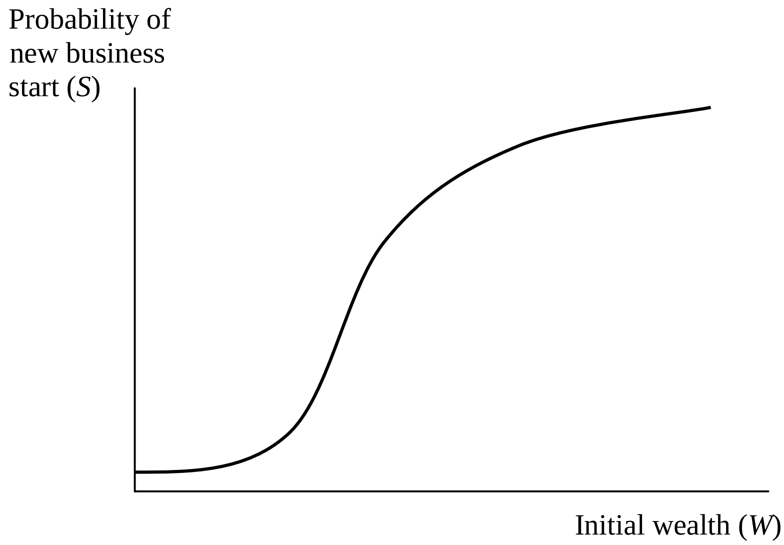
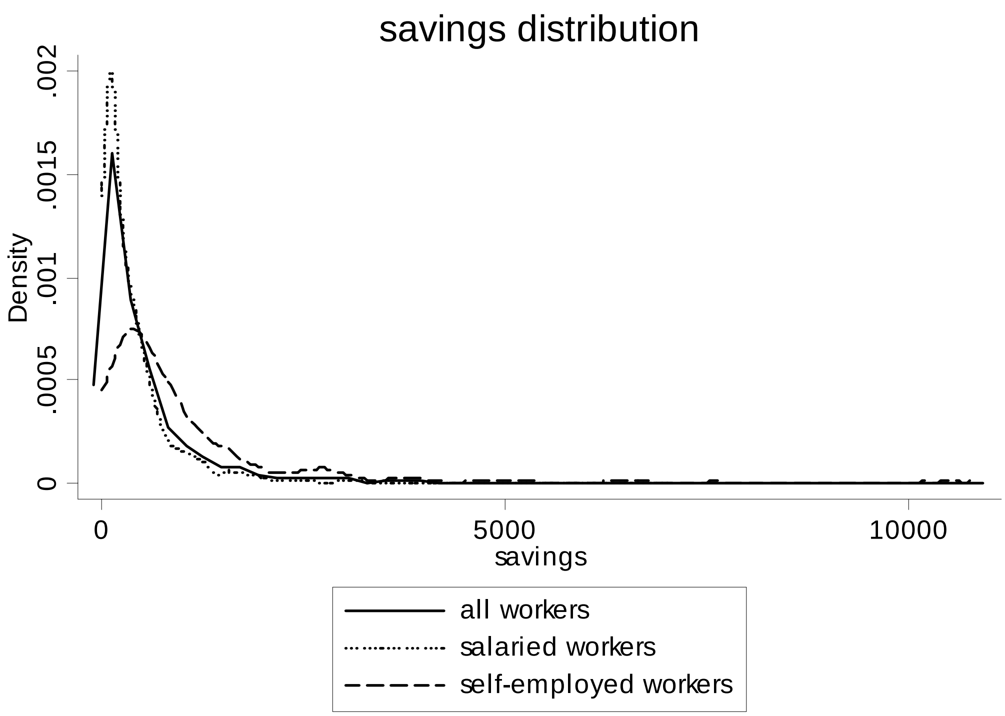


Figure 2: Densities of savings

(a) Big sample (one outlier dropped)



(b) Small sample (savings smaller than 2000 dinars)

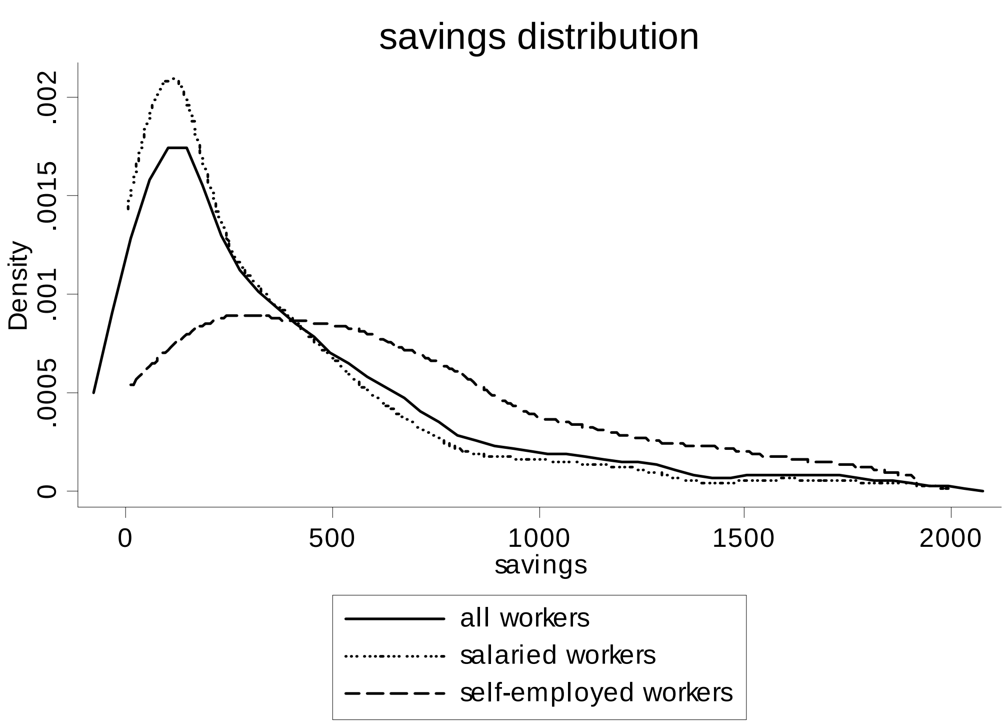
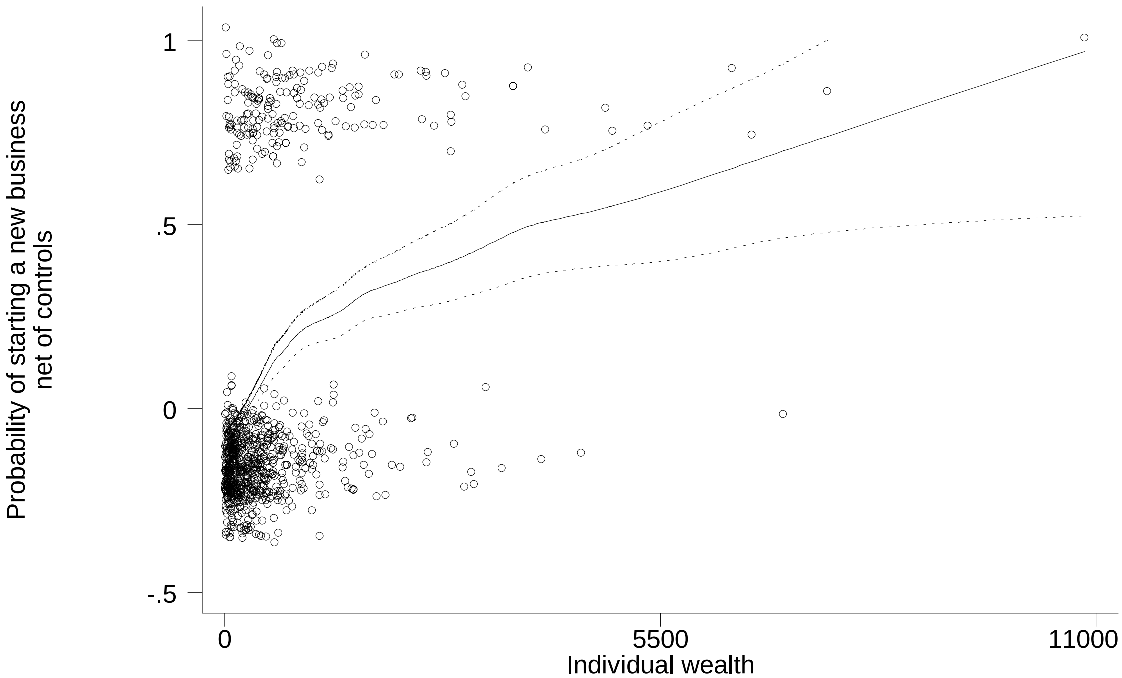


Figure 3: LOWESS regression of business starts on initial savings with and without controls (including 95% confidence intervals); sample of return migrants.

(a) With controls; dependent variable is: $S_i - X_i \hat{\pi} = \varphi(W_i)$



(b) Without controls, dependent variable is Si

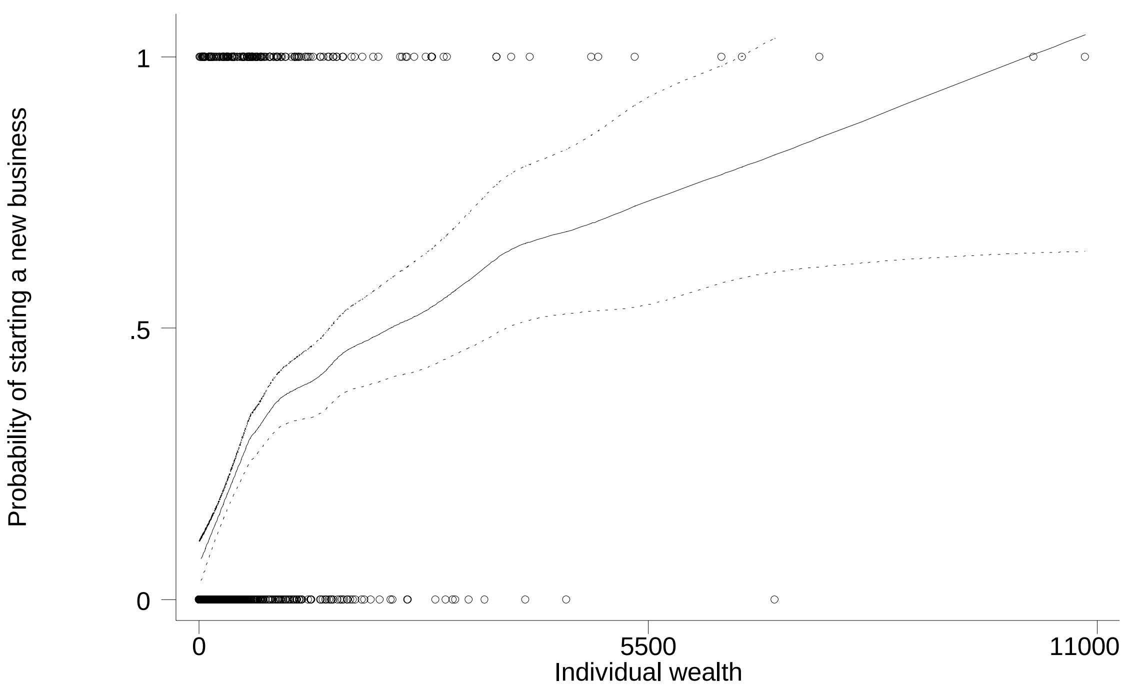
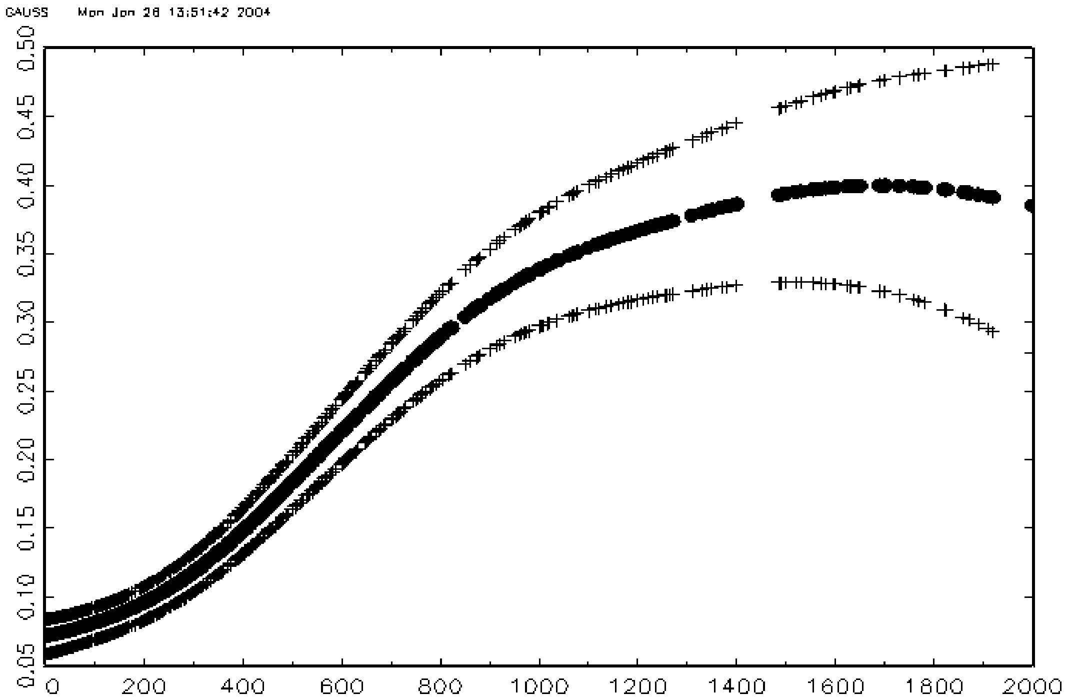


Figure 4: Non-convexity at low wealth
(the 95% poorest individuals in the sample)

(a) With controls, dependent variable is: $S_i - X_i \hat{\pi} = \varphi(W_i)$



(b) Without controls, dependent variable is S_i

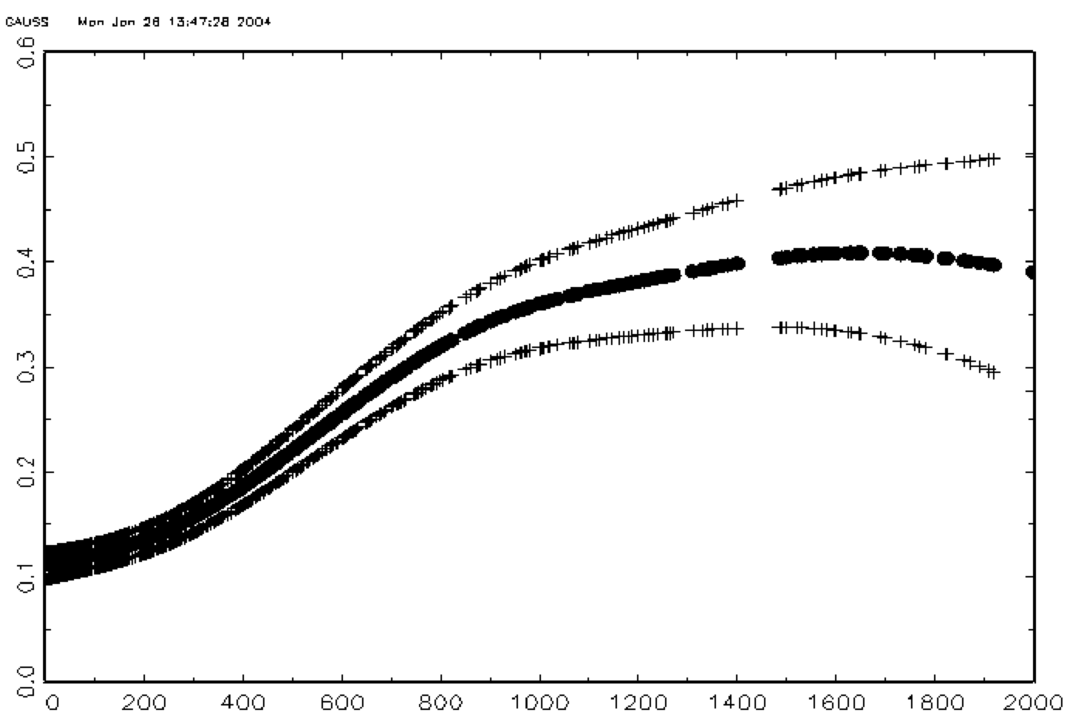


Table 1: Descriptive statistics of the two samples

Non migrants not self-employed in 1974	(1) Took up self-employment between 1974 and 1986 (n=107)	(2) Did not take up self-employment (n=499)	Full sample (n=606)
Age in 1986	37.23 (15.66)	38.15 (14.49)	37.99 (14.69)
No education (%)	29	35.1	34
Primary school (%)	52.3	31.5*	35.1
Short secondary school (%)	5.6	6.8	6.6
Long secondary school (%)	13.1	26.6*	24.3
Number of dependents	3.72 (3.46)	3.94 (3.18)	3.9 (3.23)
Married (%)	70.1	72.3	71.9
Born in area of Tunis (%)	3.8	7	6.5
Born in Center East (%)	19.8	18.3	18.5
Born in Center West (%)	22.6	20.7	21
Born in Southern East (%)	28.3	20.3	21.7
Born in Southern West (%)	7.6	12.2	11.4
Born in Northern East (%)	4.7	9	8.3
Born in Northern West (%)	13.2	12.5	12.6
Savings accumulated during 1974-1986 (1 dinar in 1986 = $US1.6)	689.04 (1464.92)	396.57* (652.40)	448.21 (859.53)

*significantly different from the mean in column (1), t-test. Standard errors in parenthesis

Return migrants not self-employed before migration	(1) Took up self-employment on return(n=210)		(2) Did not take up self-employment (n=840)		Full sample (n=1050)	
Age in 1986	38.92	(10.92)	36.75*	(11.07)	37.18	(11.07)
Age at return	34.87	(10.60)	32.44*	(10.66)	32.93	(10.69)
no education (%)	36.4		33.3		33.9	
Primary school (%)	46.4		50.2		49.5	
Short secondary school (%)	4.3		4.7		4.6	
Long secondary school (%)	12.9		11.8		12	
Number of dependents	5.0	(2.94)	4.55	(2.93)	4.64	(2.93)
Married (%)	81.4		80		80.3	
Born in area of Tunis (%)	5.3		4.9		5	
Born in Center East (%)	23.3		18.6		19.5	
Born in Center West (%)	21		24		23.4	
Born in Southern East (%)	15.7		23.1*		21.6	
Born in Southern West (%)	10		10.6		10.5	
Born in Northern East (%)	7.6		6.1		6.4	
Born in Northern West (%)	17.1		12.7		13.6	
Savings accumulated abroad (1 dinar in 1986 = $US1.6)	1086.2 (1539.13)		442.35* (951.77)		580.52 (1134.70)	

* significantly different from the mean in column (1), t-test. Standard errors in parenthesis

Table 2: First-stage regressions for accumulated savings while abroad

Instruments	Determinants of savings
age at return	27.89**
	(2.94)
square of age	-0.26**
	(-2.36)
no education	-169.6**
	(-3.12)
long sec. school level	382.2**
	(3.96)
more than long secondary	237.7**
	(3.75)
number of dependents	-8.95
	(-1.03)
Married	21.08
	(0.33)
born in area of Tunis	120.85
	(1.14)
born in Center East	-29.47
	(-0.39)
born in Center West	74.08
	(1.05)
born in Northern East	43.57
	(0.45)
born in Northern West	174.7*
	(2.19)
born in South East	-2.28
	(-0.03)
having migrated before 74	450.8**
	(8.89)
constant	-218.37
	(-1.21)
Adjusted R squared	0.1534

Number of observations: 887; t ratios in parenthesis; * significant at 5% level. **significant at 1%level

Table 3: Probit for probability of starting a business after return and exogeneity test

	dF/dx	z-stat.
Savings	0.00016**	7.72
age at return	0.0121	1.73
square of age	-0.0001	-1.37
no education	0.0106	0.28
short secondary school	-0.0783	-1.27
long secondary school	-0.0365	-0.87
Married	-0.1199**	-2.37
number of dependents	0.0083	1.40
born in Center East	0.1105*	2.00
born in Center West	-0.0093	-0.19
born in Northern East	0.0096	0.14
born in Northern West	0.0156	0.28
born in South East	-0.0556	-1.14
born in area of Tunis	-0.0234	-0.33
Log likelihood	-405.8809	
observed frequency	0.206	
predicted frequency at mean var.	0.186	
number of observations	887	
Exogeneity test:$^{\Psi}$	coefficients:	z-stat :
residuals of savings	-0.0004	-1.49
Log likelihood :	-404.7774	

Notes: *significant at 5% level; ** significant at 1% level; 1% richest individuals dropped out of the sample; dF/dx gives the effect of a small change in each continuous independent variable. For dummy variables it is equal to the discrete change in probability when the dummy variable changes from 0 to 1. $^{\Psi}$ coefficient associated to the residuals of the instrumented regression for savings imbedded into the main regression (other coefficients associated to the other explanatory variables are not reported).

Table 4: Parameters on control variables in explaining the probability of starting a business for the two samples

	Non Migrants	Return migrants
age in 1986	-.0164	
	(-1.59)	
age in 1986 squared	0.0002	
	(1.70)	
age at return		0.011
		(1.84)
age squared		-0.0001
		(-1.54)
no education	-0.083	0.053
	(-1.46)	(1.2)
short secondary school	-0.068	-0.103
	(-0.90)	(-1.68)
long secondary school	-0.177**	-0.021
	(-3.8)	(-0.43)
Married	0.017	-0.133**
	(0.28)	(-2.96)
Number of dependents	-0.006	0.003
	(-0.82)	(0.40)
born in Center East	0.154**	0.107*
	(2.53)	(1.88)
born in Center West	0.102	-0.009
	(1.82)	(-0.18)
born in Northern East	-0.007	-0.003
	(-0.1)	(-0.05)
born in Northern West	0.12	0.013
	(1.9)	(0.22)
born in South East	0.144**	-0.067
	(2.5)	(-1.28)
born in area of Tunis	0.003	-0.07
	(0.04)	(-0.94)
Constant	0.002	0.002
	(0.09)	(0.13)
Observations	463	695
R-squared	0.0708	0.0443
Tests of $\gamma = 0$	F(13,541)=0.78	F(13,865)= 0.73
	P-value=0.69	P-value=0.73

Note: Robust t-ratios in parentheses. Non migrants : 3% richest individuals are dropped from the sample. Return migrants : one outlier (savings=20550) is dropped from the sample. * denotes significant at 5% level; ** significant at 1% level.

Table 5: Probits for probability of starting a business

	Return migrants (one outlier dropped from the sample) (1)		Non-Migrants (2)		Return migrants (1% richest individuals dropped from the sample) (3)	
	dF/dx	z-stat.	dF/dx	z-stat.	dF/dx	z-stat.
savings$^{\alpha}$	0.0004**	6.24	0.0008**	3.58	0.0004**	7.13
savings squared	-1.13e-07**	-3.53	-1.09e-06**	-3.08	-7.04e-08**	-4.40
savings cubic	9.80e-12**	2.68	3.90e-10**	2.73		
age at the date of survey			-0.0107	-1.42		
age (date survey) squared			0.0001	1.42		
age at return	0.0115	1.67			0.0115	1.69
age at return squared	-0.0001	-1.34			-0.0001	-1.38
no education	0.0143	0.38	-0.0874*	-2.08	0.013	0.35
short secondary school	-0.0914	-1.55	-0.076	-1.46	-0.068	-1.13
long secondary school	-0.042	-1	-0.1463**	-4.19	-0.041	-1.00
Married	-0.1278**	-2.48	0.0071	0.15	-0.1247**	-2.47
number of dependents	0.008	1.34	-0.0037	-0.56	0.008	1.35
born in Center East	0.1315**	2.3	0.145*	2.08	0.1332**	2.36
born in Center West	-0.0079	-0.16	0.1077	1.65	-0.0001	-0.00
born in Northern East	0.0131	0.19	0.0189	-0.26	0.0156	0.23
born in Northern West	0.0321	0.56	0.1476	1.93	0.032	0.57
born in South East	-0.0509	-1.02	0.1587**	2.36	-0.048	-0.98
born in area of Tunis	-0.033	-0.46	0.0273	0.32	-0.0314	-0.45
Log likelihood	-398.6894		-234.7732		-395.7985	
Observed frequency	0.2123		0.1672		0.2063	
Predicted frequency at mean var.	0.1852		0.1447		0.1782	
number of observations	895		574		887	

Notes: *significant at 5% level; ** significant at 1% level ; model estimated between 1974 and 1986 for non migrants (3% richest individuals dropped out of the sample) and after return for return migrants. dF/dx gives the effect of a small change in each continuous independent variable. For dummy variables it is equal to the discrete change in probability when the dummy variable changes from 0 to 1. $^{\alpha}$ accumulated savings during 1974-1986 for non migrants; accumulated savings abroad for return migrants.

Table 6: Probit for probability of starting a business after return

Marginal effects: dF/dx	(1) 95% poorest individuals	(2)	(3)	(4)
Savings	0.0005	0.0002	0.0003	0.0002
	(5.61)**	(6.24)**	(7.91)**	(7.36)**
savings squared	-0.0002	-1.15e-08	-2.13e-08	-1.63e-08
	(3.36)**	(2.07)*	(4.65)**	(3.80)**
age at return	0.0108	0.006	0.019	0.013
	(1.66)	(0.84)	(2.23)*	(1.81)
age squared	-0.0001	-0.00004	-0.0002	-0.0001
	(1.36)	(0.46)	(1.69)	(1.45)
no education	0.0104	0.042	-0.018	0.006
	(0.29)	(1.05)	(0.42)	(0.15)
short secondary	-0.0702	-0.129	-0.130	-0.067
	(1.16)	(1.86)	(1.89)	(1.04)
long secondary	-0.0520	-0.055	-0.046	-0.033
	(1.29)	(1.14)	(0.95)	(0.75)
born Center-east	0.1451	0.191	0.099	0.141
	(2.59)**	(2.92)**	(1.56)	(2.41)*
born Center-west	0.0067	-0.016	-0.058	-0.006
	(0.14)	(0.31)	(1.08)	(0.12)
born Northeast	0.0004	-0.017	-0.063	-0.028
	(0.01)	(0.23)	(0.84)	(0.38)
born Northwest	0.0421	0.009	-0.035	0.015
	(0.62)	(0.13)	(0.50)	(0.22)
born Southeast	0.0206	0.033	0.017	0.034
	(0.37)	(0.54)	(0.26)	(0.57)
born Tunis area	-0.0263	-0.026	-0.083	-0.033
	(0.54)	(0.48)	(1.52)	(0.62)
married	-0.1201	-0.138	-0.182	-0.139
	(2.40)*	(2.39)*	(3.02)**	(2.51)*
No. dependents	0.0099	0.011	0.010	0.009
	(1.74)	(1.63)	(1.41)	(1.48)
skilled pre-migration		0.051		
		(1.61)		
skilled post return		0.114		
		(2.06)*		
unemployed pre migration		-0.004		
		(0.11)		
duration since return				0.004
				(0.97)
unemployment duration at return			-0.007	
			(3.04)**	
Observations	854	742	754	863

Notes : Absolute value of z statistics in parentheses. * significant at 5%; ** significant at 1%

Table 7: Controlling for heterogeneity correlated to migration country

Marginal effects : dF/dx	(1)	(2)	(3)	(4)
	Adding controls for migration country	Adding control for Libya	Migrants from Libya only	Migrants from Libya only; savings <2000 dinars
Savings	0.0002	0.0002	0.0003	0.0005
	(6.63)**	(6.79)**	(4.71)**	(4.64)**
Savings squared	-1.48e-08	-1.49e-08	-3.95e-08	-1.85e-07
	(3.54)**	(3.57)**	(1.78)	(2.78)**
age at return	0.014	0.012	0.012	0.010
	(1.95)	(1.78)	(1.24)	(1.05)
age squared	-0.0001	-0.0001	-0.0001	-0.0001
	(1.67)	(1.49)	(1.04)	(0.88)
no education	0.014	0.016	0.018	0.018
	(0.38)	(0.43)	(0.46)	(0.48)
short secondary	-0.096	-0.094	-0.103	-0.098
	(1.62)	(1.57)	(1.59)	(1.60)
long secondary	-0.036	-0.045	0.001	-0.008
	(0.85)	(1.09)	(0.02)	(0.17)
born Center-east	0.134	0.127	0.213	0.204
	(2.34)*	(2.24)*	(3.27)**	(3.23)**
born Center-west	-0.004	-0.011	0.013	0.005
	(0.09)	(0.22)	(0.24)	(0.10)
born Northeast	-0.036	-0.038	0.007	0.015
	(0.51)	(0.53)	(0.08)	(0.18)
born Northwest	0.006	0.005	0.084	0.100
	(0.09)	(0.08)	(1.03)	(1.23)
born Southeast	0.020	0.016	0.042	0.021
	(0.36)	(0.28)	(0.65)	(0.34)
born Tunis area	-0.046	-0.051	0.010	0.006
	(0.92)	(1.01)	(0.17)	(0.12)
Married	-0.144	-0.130	-0.131	-0.122
	(2.77)**	(2.53)*	(2.47)*	(2.35)*
No. dependents	0.008	0.008	0.005	0.004
	(1.28)	(1.31)	(0.74)	(0.76)
France	0.197			
	(2.16)*			
Libya	0.056	-0.088		
	(0.80)	(2.52)*		
Europe	0.220			
	(1.74)			
Observations	895	895	684	672

Notes: Absolute value of z statistics in parentheses; * significant at 5%; ** significant at 1%; in column (1) the missing category for countries of migration is other Arabic countries