Heterogeneity among Mexico's Micro-Enterprises

An Application of Factor and Cluster Analysis

Wendy V. Cunningham

William F. Maloney

The small firm sector comprises several types. Although some are a refuge for displaced salaried workers, most were started by workers who chose to be entrepreneurs — and they show dynamics consistent with patterns in the industrial world.

Summary findings

A long tradition sees the small firm sector as a holding pattern for workers queuing for jobs in the formal sector of a segmented labor market. An alternative “entrepreneurial” view suggests that many workers prefer self-employment to salaried jobs. These competing views can be resolved if the sector is heterogeneous.

Using factor and cluster analysis, Cunningham and Maloney generate a typology of the sector by taking advantage of a Mexican data set on micro-firms that offers information on a broad range of small firm characteristics. The methodology permits divisions to emerge from the data without the a priori imposition of a theoretical structure.

The data break into several distinct groups, broadly characterized as highly profitable and dynamic young firms, older firms that have stabilized at a small size, and young firms that act as an employer of last resort. Those in the last group, comprised of older entrepreneurs with low levels of education, are the most likely to cite that they started their firms because they were unable to find a salaried job.

In general most of the firm owners in all groups stated that they chose self-employment over formal sector

employment in order to be independent, collect higher earnings, or follow family tradition. These survey responses are supported by the finding that income distribution adjusted for human capital is composed of two subdistributions, with the “underperforming” distribution comprising only 14 percent of the sample.

The factor analysis also implies that firm owner characteristics and firm size or profitability may not be correlated. For example, young workers who we might think are forced into the small firm sector due to inability to enter the formal job market do not necessarily earn less or have less capital than older entrepreneurs. Furthermore, a distribution of the earnings residual factor shows that very few firms, regardless of the firm owner’s age, are earning below their expected profits.

The data suggest that the small size of informal firms may not necessarily result from limited access to financial institutions or a desire to evade labor or tax laws. Instead, the firms simply may be in the beginning stages of a growth process or owners may prefer to remain small.

This paper is a product of the Poverty Reduction and Economic Management Sector Unit, Latin America and the Caribbean Region. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Tania Gomez, room I8-102, telephone 202-473-2127, fax 202-522-2119, Internet address tgomez@worldbank.org. William Maloney may be contacted at wmaloney@worldbank.org. October 1998. (42 pages)

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.

**Heterogeneity among Mexico's Micro-Enterprises:
An Application of Factor and Cluster Analysis***

Wendy V. Cunningham

William F. Maloney

* We thank the Mexican National Institute of Statistics, Geography, and Information (INEGI) for the use of the data. INEGI is in no way responsible for any incorrect manipulation of the data or erroneous conclusions drawn from it.

I. Introduction

Owners of small firms, defined as those of under six employees, constitute 20-25% of the Mexican work force and those working for these firms roughly 10%. Both their importance as sources of employment and their tendency to be found in the lower end of the income spectrum makes understanding the dynamics and reason for being of these firms essential to both issues of labor market efficiency and poverty and distribution.

A long tradition views the sector as the subsistence holding pattern for workers queuing for jobs in the formal sector of a labor market segmented by above market clearing labor costs.² However, an alternate view dating from Keith Hart's (1972) accounts of the dynamism of Kenya's micro-entrepreneurs takes the basic unit of analysis as the *firm*, rather than the disadvantaged worker, and stresses the large overlaps in the distribution of remuneration between the self-employed and salaried sectors and that, as in the industrialized world, workers often prefer self-employment to salaried jobs.³ Within this firm-centered view coexist further contradictory hypotheses on firm dynamics and the reasons that small firms are small, such as satisficing behavior vs. structural constraints to growth or a desire to evade formal institutions vs. lack of access to them.

These conflicting views can be reconciled if, in fact, the sector is very heterogeneous and contains elements of each story, a possibility acknowledged by the literature in its increasing reference to "upper" and "lower" tiers. However, to date, the data on the micro-firm sector has been largely anecdotal and offers little evidence for this division, or other divisions of equal

² The Harris-Todaro (1970) model is perhaps the traditional statement of this view. Rigidities or excess remuneration may arise from severance pay, social security, medical benefits, high minimum wages, and costly regulation in general. See Stiglitz(1974) and Esfahani and Salehi-Isfahani (1989) for an efficiency wage approach.

³ See Tokman (1989), Fields (1990), Portes and Shauffler (1992).

interest. This paper takes advantage of a large and comprehensive data set on Mexican micro-firms that offers information on a broad cross section of small firm characteristics along which we may find division or segmentation and that makes generating a richer typology feasible. To this end we employ two econometric techniques, factor and cluster analysis, that permit the segmentation of the market to emerge from the data with a minimum of prior structure imposed. These techniques have been employed in the literature on industrial or labor market dualism in the industrialized countries by Oster (1979) and Anderson et. al. (1987) for the US, Sloane et. al. (1993) for the UK, and Flatau and Lewis (1993) for Australia.

II. Possible Sources of Heterogeneity

In the spirit of Hart, Levenson and Maloney (1997) argue that small firm dynamics in LDCs may be well described by models now standard in the industrialized country literature rather than representing qualitatively different phenomena. These models imply heterogeneity that is unrelated to structural features of LDCs or labor market distortions. Lucas (1978) argued that the vast range of firm size among established firms can be derived from a model where entrepreneurs have differing cost structures. Building on this, Jovanovic (1982) offers a model of firm dynamics where entrepreneurs face a multiplier on their costs, $x(\theta+\epsilon)$, that is a function of entrepreneurial ability or perhaps firm location, θ , and ϵ captures random noise:

$$C = x(\theta+\epsilon)c(q). \quad (1)$$

It is assumed that $c'(q)>0$ and $c''(q)<0$ where q is the level of output. Entrepreneurs have only a vague idea of what their true cost structure will be at the time of start up and only with experience can they generate a more precise estimate. If they find themselves with profits above

those expected they will lower their estimates of their costs, invest, and expand to their long run size. Others may find themselves far less profitable than expected and go out of business. This framework has two implications.

First, the observed distribution of mature firm size and the co-existence of larger, more capital intensive, high earnings firms with very small ones may only reflect the underlying distribution of whatever factors contribute to θ . Assertions of the prevalence of “pre-capitalist,” “non-optimizing,” or “satisficing” modes of production (Tokman 1992 and Elizondo 1992) may be nested if we argue that entrepreneurs are maximizing if not profits, then utility more generally and that the high implicit valuation of additional labor after basic needs are satisfied makes θ very high. Second, among small firms we should find long lived firms at their long run size, start-up firms that find themselves extremely profitable and desire to expand, and start-up firms that are about to fail. We should also find differing returns to factors that reflect this normal heterogeneity. On average, the youngest firms should show the highest mortality rates since they have the least refined estimates of their viability and are more likely to discover that they are not than more established firms.⁴ In sum, small size, lack of growth, or even differential returns to human and physical capital may be due to variance in individual abilities and preferences and the noisy process of discovering them instead of due to distorted labor and credit markets or an inefficient regulatory framework, which themselves may give rise to other divisions in the small firm sector.

Labor market distortions that lead to a rationing of workers from their preferred position in salaried work should lead to segmentation in the micro-enterprise sector as well. First, we

⁴ Numerous US and UK surveys support these models showing extremely high rates of firm mortality among young and small firms that decrease with age and size (See Davis et al 1994, Evans and Leighton 1989, Dunne et al 1989).

should find concentrations of young workers, migrants, and perhaps laid off middle age workers who report having trouble finding salaried work coexisting with those who voluntarily became self-employed. Further, we should find the standard differences in earnings equations postulated in the dual labor market literature. Returns to the human capital of those voluntarily in the sector should be comparable to those in salaried employment while those involuntarily in the sector who were rationed out of "good" jobs would be lower.⁵

Credit constraints on entrepreneurs are evident in the industrialized countries (Evans and Jovanovic 1989) and are very likely in LDCs as well. "Life cycle" behavior where workers enter into salaried work; accumulate knowledge, capital, and contacts; and then quit to open their own businesses may represent a natural life progression, but may also reflect credit market imperfections, or even the poor quality of the education system. If credit market imperfections constitute an important barrier to growth for some firms, we should expect to find a group with low θ 's, earning supernormal profits, who desire to expand, but lack access to resources.

There is no consensus in the literature on how compliance with labor or tax legislation is related to production decisions. To avoid taxes or regulatory legislation, firms may choose to stay small or if they grow large, the firms may subcontract to smaller firms (Loayza 1997, Rauch 1991, Portes and Shauffler 1993). But, de Soto (1989) asserts that many aspects of "formality" are essential for the growth and efficient operation of firms and are something for which firms are willing to pay, but are often denied by oppressive bureaucracies. Nesting both, Levenson and Maloney (1996) argue that participation in the institutions of civil society - for example those

⁵ In fact, if the return to factors in the formal sector were imputed, many involuntarily self-employed might have cost structures that would dictate that they close. In the absence of unemployment insurance, a laid off middle aged worker unable to find salaried work may invest his severance pay to generate employment, perhaps buy a taxi, even

that grant secure property rights, permit formal contracting mechanisms, or access to formal financing mechanisms - can be considered a factor of production that becomes increasingly important as firms grow and that has a cost measured in terms of compliance with legal norms. For many high θ , and therefore permanently small firms, issues of formality may therefore be largely irrelevant .

III. Data

The data are from the 1992 National Micro-Enterprise Survey that was collected by the Mexican National Statistical Institute in 1992:1. This survey was constructed by selecting 11,000 owners of micro-enterprises, defined as firms of under six individuals (16 for manufacturing firms), from the last quarter of 1991 of the National Urban Employment Survey and examines reasons for becoming self-employed, incomes, capital structures, costs, enterprise problems, credit needs, migrant status, and employment patterns, among others. To focus on the firms commonly thought of as informal, we restrict the sample to males ages 16 - 65 with no more than 12 years of formal education.

The analysis include four broad classes of variables, used both raw and transformed into indices (see appendix for more detail):

A. Characteristics of the Entrepreneur:

1. *Years of Education of the Entrepreneur* (Education).
2. *Years of Work Experience of the Entrepreneur* (Experience).
3. *Why initiated enterprise* (Voluntary). Captures whether the entrepreneur is voluntarily employed in the sector.

though if the cost of capital were fully imputed, the firm might not be viable. The worker is borrowing against future (past) income to subsidize an unviable firm.

B. Characteristics of the firm:

1. *Firm Life in Years* (Firm Life).
2. *Capital-Labor Ratio* (K/L). Reported total capital stock of the firm divided by the total number of workers in the firm.
3. *Earnings* (Earnings). Monthly total revenues of the firm less expenses.
4. *Permanence of Work site* (Site). Whether the firm has a permanent site or not.
5. *Average Hours Worked by Entrepreneur per Week* (Hours).
6. *Total number of workers in the firm* (Workers).

C. Entrepreneurial Dynamics:

Since the sample does not permit explicitly testing for dynamics, we rely on the entrepreneur's statements of his plans and needs for financing.

1. *Permanence in the Sector* (Move/stay). Whether the entrepreneur plans to close the firm and look for salaried work.
2. *Plans to Expand* (Expand). Whether the owner plans to go out of business, remain in business but not expand, or expand.
3. *Clients* (Clients). captures the nature of clientele ranging from local families to large businesses.
4. *Initial financing for firm* (Financing-Initial). captures the source of start up capital, ranges from personal savings, friends, and informal markets up to formal banking institutions.
5. *Demand for additional financing* (Financing-New): Whether the firm has sought additional financing from any source.
6. *Financing problems* (No Credit Problems). Whether the owner cites credit availability as a business problem.
7. *No problems* (No Problems). Whether the firm reports having no business problems.

D. Participation in formal market and legal institutions

1. *Registration with Tax Authorities* (Tax). Ranges from being unregistered with tax authorities, to being registered locally, to being fully registered with federal tax authorities.
2. *Compliance with Labor Legislation* (Benefits). What fraction of workers are covered by social security and health benefits.

IV. Empirical Methodology

*i. Factor Analysis*⁶

Factor Analysis is concerned with finding a small number of common factors that linearly reconstruct a large number of variables:

$$z_{ij} = \sum_{p=1}^h a_{jp} F_{pi} + e_{ij} \quad (2)$$

Where z_{ij} is the value of the i^{th} observation on the j^{th} variable, F_{pi} is the i^{th} observation on the p^{th} common factor, a_{jp} is the weighting on that factor, and e_{ij} is known as the j^{th} variable's unique factor and is analogous to the residual in standard regression analysis. The factors themselves are linear combinations of the variables such that:

$$F_{pi} = \sum_{j=1}^k q_{pj} z_{ji} \quad (3)$$

Where F_{pi} is the value of factor p where there are h selected factors. for individual i for each of the n individuals with observations on k variables and q is the weighting of the p^{th} factor in variable j . By isolating a subset of the k possible factors, a large and seemingly chaotic mass of data can be found to be structured along a few critical dimensions.⁷

⁶ See Harman (1976) for a detailed review of the technique and Kim and Mueller (1978) for a more general review.

⁷ The initial data matrix is rotated to generate a diagonal matrix where the eigenvalues reveal how much of the variation occurs along each dimension. A subset with the highest explanatory power are chosen and the resulting

The technique has four attractive features that aid in teasing out the composition and dynamics of the micro-firm sector. First, it isolates covarying variables that suggest types of entrepreneurs or firms, that would not be possible from simple bi-variate cross correlations, again, without imposing any prior structure on the data. Second, if in fact there are “good” and “bad” firms within the micro-firm sector, we should be able to identify a “dualistic” factor that loads heavily on the characteristics commonly associated with firm “quality.” Third, since the value of any particular factor differs for each observation, the factor will show a distribution that offers a statistical test for segmentation. If, in fact, there are two types of firms with distinctive modes of operation and dynamics, we might expect the observed distribution to be comprised of two underlying distributions corresponding to the inferior and superior sub-sectors:

$$F \sim p N(\mu_1, \sigma_1^2) + (1-p) N(\mu_2, \sigma_2^2) \quad (4)$$

Using the EM algorithm, we maximize the likelihood across both means, standard deviations and the weighting on each distribution, the share of the population found in each segment, p , $(1-p)$.

We then compare the likelihoods to test against the null that the data is drawn from a single normal distribution. Fourth, because the factors are mutually orthogonal by construction, they cannot be causing the segmentation or patterns observed in each other.⁸

matrix is rotated again to aid in interpretation by re-expressing the factors so that the loadings on a few variables are as large as possible. The principal component method of extracting factors was employed rather than the common maximum likelihood or other methods because it does not require normality in the distribution of the variables. Though the various alternative rotation methods yielded similar results, the Varimax method offered the clearest separation of factors and preserved the orthogonality of the initial factors.

⁸Since the factors identified may be sensitive to the variables included in the analysis, we ran several specifications. In particular, variables with very high loadings were eliminated to see if that variable alone was driving the factor. In general, the same factors formed with the same economic interpretations.

ii. Cluster Analysis ⁹

The Ward clustering method combines observations on entrepreneurs and their firms into progressively larger endogenously determined clusters by minimizing the sum of the within group variance of all clusters:¹⁰

$$W = \sum_{k=1}^{k=g} \sum_{j=1}^{j=m} \sum_{i=1}^{i=n_k} (x_{ijk} - \bar{x}_{jk})^2 \quad (5)$$

where $\bar{x}_{jk}$ is the mean value of variable j in cluster k , x_{ijk} is the value of an observation assigned to cluster k , n_k is the number of observations in cluster k , m is the number of variables, and g is the number of clusters. Starting with n clusters each consisting of a single observation, this Ward value is calculated n times and takes a value of 0, when $n=k$ and the distance of the n observations from the n cluster centers (themselves) is zero, and then increases as more distant points are forced to share a common center until $n=1$. The optimal number g of clusters is decided and each observation is assigned to the nearest centroid, or mean vector.¹¹

⁹ See Anderberg (1973), Everitt (1974) and Späth (1980) for more detailed treatments.

¹⁰ Cluster analysis requires some prior manipulation of the data. First, it is very sensitive to magnitudes, so each variable is normalized to a scale with a range of [0,1]. Second, extreme values or outliers compress the bulk of the observations into a very limited range and yield trivial clustering results. In some cases, therefore, these values are suppressed or recoded to provide a greater variance across the majority of the sample.

¹¹ The n values are used in two ways to determine the number of clusters. First, a dendrogram graphically depicts the hierarchical combination process as clusters are formed and consolidated which suggests the appropriate level of consolidation. A second, but related method is to stop clustering when the Ward value rises above a certain threshold or the second derivative of the plot suddenly increases. This reveals when agglomerations are relatively more “forced” and where the data should be left in distinct sub-clusters. Observations are assigned using the standard MacQueen's k-Means Method. Since the resulting clusters may be sensitive to the initial centroids, the analysis was run using different beginning seeds for both steps. Regardless of the initial seed used, the final centroids were very similar and the characterization of the clusters did not change.

V. Results

i. Continuous variables

Although in both types of analysis, any kind of variable can be included, in practice, discrete or non-continuous variables can introduce trivial bi- or multi-modality in the distributions of factors, and their extreme values tend to dominate the clustering process. We therefore begin working with just the six continuous variables, education, experience, firm life, weekly hours worked, capital per worker and earnings. In various combinations, these variables emerge as elements of segmentation stories: marginal, under-employed workers in low profit, low capital firms vs. better capitalized prosperous firms; young workers unable to enter formal employment in short lived firms vs. established voluntary entrepreneurs. If such discrete subgroups exists, both techniques should identify them. We then draw on more descriptive evidence on motivations for being in the sector, business problems, degrees of compliance with social norms and test directly for segmentation among these groups by comparing estimated earnings equations. Drawing together these sources, we generate a typology of the subgroups.

ia. Factor Analysis

The loadings on two factors that explain 56% of the variance observed in the data are presented in table 1a.

Factor 1: Generational characteristics (32% of the variance). The factor loadings show firm life and worker experience entering strongly positively and education strongly negatively. Younger workers are better educated and, clearly, cannot have been in business as long as their older counterparts.

Factor 2: "Dualism?" (23% of the variance) What might be considered as a dualism factor loads heavily and positively on earnings, hours worked by the entrepreneur, capital-labor ratio, and to a lesser degree, education and firm life.

The orthogonality of these two factors implies that there is no correlation between generational effects and those determining capital intensity or prosperity. Of greater interest, however, is the evidence for dual distributions in each factor. Table 1 suggest that the distributions sketched in figure 1 are mixed. The first factor contains a distribution comprising roughly 43% of the sample that has higher levels of education, relatively less experience, and, logically, shorter firm life, and a second with lower levels of education and more experience.

For the tentative dualism factor (figure 1b) 26% of the population corresponds to a distinct distribution centered at a higher level of capital per worker, earnings, hours worked, firm life, and education, thus providing some support for an upper tier/lower tier division. However, dualism in the sense of different returns being earned for the same productive factors can only be tested by adjusting earnings for these endowments. We employ a standard earnings equation augmented by returns to capital and interpret the residual, η , as a measure of earnings above or below what would be expected given levels of human and physical capital:¹²

$$\text{Ln(Earnings)} = \beta_0 + \beta_1 \text{Educ} + \beta_2 \text{Exp} + \beta_3 \text{Exp}^2 + \beta_4 \text{Exp*Edu} + \beta_5 \text{K} + \beta_6 \text{Ln(Hours)} + \eta \quad (6)$$

¹² Taking the log of hours is standard in wage equations because when β_6 is constrained to be 1, the dependent variable becomes the log of hourly wages. Its inclusion here is somewhat problematic since returns to capital should not be a function of hours worked necessarily. Though not optimal, the present specification at least lets the data identify the value of the coefficient.

Table 2 presents the results of replacing earnings with the residuals. Factor 1 again captures a generational effect. A “capital intensity” factor emerges that suggests that entrepreneurs with more capital per worker tend to work longer hours, tend to be better educated and, consistent with Jovanovic’s approach, have been in business longer.

Factor 3 weighs heavily on the residual, but on none of the other variables important in the previous “dualism” factor. Further, the MLE estimates suggest only 13% of the sample is found in a distinct, lower distribution that might be attributable to the disadvantaged sector of a dual labor market. The positive association of earnings in the “dualism” factor arises from the distribution of human and physical capital rather than differing returns to them.

ib. Cluster Analysis

The results from the cluster analysis tell a broadly similar story. The mean values for the six clusters are given in table 3, the distance between the final cluster centers in table 4, and the mean values converted back into their original range in table 5.¹³ Tables 6-8 provide information on a variety of firm and enterprise characteristics by cluster.

The generational division identified in factor 1 broadly emerges, with clusters 1-3 being largely young and capturing 53% of the sample and 4, 5 and 6 capturing older and less well-educated workers. The other divisions correspond broadly to the second factor which, as the orthogonality of the factors would predict, cuts across generations. Clusters 2 and 4 are distinctly higher in capital intensity, although the high earnings, hours worked, and education suggest that cluster 1 could also be part of an upper tier comprising 25% of the sample-- close to the 26% suggested by the MLE.

¹³ In reality, the dendrograms leaned slightly more heavily to there being five factors but the six factor break, while similar, yielded an additional important division.

Though we would again like to compensate for human and physical capital, running the cluster analysis using the residuals from the earnings equation proved problematic.¹⁴ As an alternative, table 9 presents the results of the earnings equations for each cluster. First, the returns to education, experience, capital, hours worked and other coefficients were constrained equal leaving just the intercept to vary. Given the sample-wide return to those factors, this tells how much each sector earns above or below that of the most numerous case, cluster 3. The specification accounts for 33% of the variance and the principal coefficients are significant and of the predicted sign with positive returns to education, experience, capital, and hours worked. The next 6 columns allow the returns to factors to vary as well. The asterisks on the individual coefficients, and the joint tests on the coefficients taken as a group indicate statistically significant differences from cluster 3.

ic. A Tentative Typology Using Continuous Variables

Cluster 1: Very prosperous, low capital-intensity firms (14%). Entrepreneurs in this sector revealed very low involuntary entrance (17.0%) and very high desires for independence, higher pay, and continuing family tradition. With a mean age of 39 years, 78% were salaried employees before. They have very little capital, but are the largest firms (1.02 employees), and earn roughly 2.5 times the lowest sector. The very high mean wages, even when controlling for education and experience, suggest that they may have informal or on the job training or skills not captured by the education proxy. Unsurprisingly, they report the fewest problems with low earnings or excessive competition.¹⁵ They are among the least likely to exit and the most likely to

¹⁴ The relatively long tails of the residuals distribution led to cluster means that showed virtually no differences in adjusted returns. While truncation of the tails yielded enough “spread” in the variable to generate some variance, the results were somewhat sensitive to how it was done.

¹⁵ The vagueness of the motivations questionnaire reported in table 7 makes interpretation of the exact nature of several of the business problems very difficult. In a healthy competitive environment, complaints of “excessive

expand. Although they relied primarily on their own resources for start up and have little demand for additional financing, they do appear relatively more constrained for credit than all but cluster 2. This sector is clearly the best off of all the sectors, although, importantly, it is not the sector with the most capital per worker.

Clusters 2 & 4: Highly capitalized firms (7% and 4%) Paradoxically, both clusters have much higher capital-labor ratios but lower earnings than cluster 1. This is likely due to the fact that both groups show 40% of their members working out of vehicles, 10 times that of the other sectors, a fact that probably explains their disproportionate problems with (traffic) authorities as well. The fact that they receive the lowest return to capital reflects perhaps over-capitalization due to the lumpiness of vehicle purchases, or that they may not fully own the vehicle. Neither shows significant returns to education, and both show extremely low returns to experience as might be expected in a vehicle operator.

It is in the demographics and the reasons for being in the sector that the two clusters differ. Cluster 2 resembles the younger cluster 1 although the percentage of entrepreneurs who left their previous job due to plant closings or being laid off (27%), and having previously worked for firms of over 50 people (43%), are nearly double that of any other sector. It is possible that they used their severance pay to invest in a car or truck. Yet the number citing inability to find another salaried job as the reason to start the business is relatively low (10.7%), and the share planning to expand the business, 30%, is as high as the prosperous cluster 1. Whatever their motivation for entering, once in, they choose to stay.

competition” are probably not uncommon. “Lack of economic resources,” “low earnings” and “lack of clients” may be covering similar ground and may not represent evidence of any distortions.

Cluster 4 entrepreneurs are clearly of the older cohort with a mean age 15 years higher than that of cluster 2, and show the longest firm life after cluster 6. As with clusters 5 and 6, being older appears to make finding alternate salaried employment less easy, but among the older groups, they are the least pessimistic and exceed all other groups in reporting the desire to be independent as the reason for leaving their previous job (59%) and starting this business (60%), or family tradition (14%). Of those who left their previous job, they have the lowest incidence (15.6%) of having been involuntarily separated. Importantly, this group is overwhelmingly voluntary in the sector but complaints about a lack of clients or low earnings are not particularly lower than other groups.

Cluster 3: Younger entrepreneurs/ startups (32%). This is the largest of the subgroups, with the youngest workers (31 years), the firms with the shortest life (3.75 years), the second lowest level of capitalization, and earnings adjusted for sample wide returns to human and physical capital between 10 to 86% less than those of every other group (table 9). These workers show the second highest levels of pessimism about finding other salaried work, (15.2%), downplay independence as a motivation (35.1%), and cite complementing family income as an important reason for starting the business (14%).

But there are many reasons *not* to think of the cluster as the safety net lower tier resulting from a segmented labor market, but rather as the normal start up firm sector with its very high mortality rates, and a higher proportion of entrepreneurs making insufficient or negative profits. First, involuntary entries still constitute only 20% of the total, and 32% said they left their previous job because it did not pay enough, the highest of any cluster. Many may be young workers, starting out, and shopping among sectors for good jobs. While it is true that 64% report having insufficient clients as their principal business problem, the highest of any group, this

might be expected of younger firms establishing their reputation, and even the clearly voluntary and prosperous sectors 1 and 4 report rates around 55%. They do not complain disproportionately about excess competition or low earnings, nor is their growth particularly constrained by non-availability of credit. Fewer do plan to expand and more plan to leave business, but again, this may simply reflect the fact that a large fraction of younger firms may go out of business.

Cluster 5: Life-cycle entrants, mid-life lay-offs, and retirees (25%). This group comprises a quarter of the sample and consists of older entrepreneurs who opened their firms relatively recently. The timing at mid-life offers support for theories of life cycle behavior, but, a relatively high percentage was laid off from their previous job and 22%, more than any other sector, report that they started this business because they could not find another salaried job. This would be the case if employers tend not to hire older workers. Double the percentage of any other group (3.7%) are retirees, and almost 8%, vastly higher than any other group, left their previous job because of illness or injury. There is little family tradition, and more than even the young, many workers (16%) started the business to augment their family's income. Adjusting for sample wide returns to human and physical capital, the group is perhaps 10% better off than base cluster 3, making it the second least well off of any sector. While the sector clearly serves as a safety net for displaced older workers, the sick and the retired, it also shows the highest rate of return to experience and the second highest rate of return to additional hours worked.

Cluster 6: Older owners of established firms with little capital (18%). These entrepreneurs are among the oldest with a mean age of 51 and firm lives of roughly 23 years. They have few plans to expand and have clearly stabilized at a small size. Yet, overall, the group is more similar to the prosperous clusters 1 and 4 than either 3 or 5. First, adjusted for sample

wide returns to human and physical capital, earnings are clearly closer to those of 2, 4, and perhaps 1, than 3 or 5. For unclear reasons, returns to their 4 years of education are negative and to experience are insignificant although they show the highest returns to capital. The cluster is second lowest in entrepreneurs entering because of lay offs or plant closings and, despite the age of these workers, the pessimism about finding a salaried job is not especially high. It holds an intermediate position both toward independence and higher pay as motivations for entering the sector. As with clusters 1 and 4, a sizable group (13.3%) started the firm because of family tradition. With the exception of reporting a relative lack of clients, its business problems are not so different from those of the prosperous sector 4.

In sum, the most natural grouping of firms by characteristics frequently associated in popular theories with segmentation do not generate sectors that are clearly the repository of the involuntarily self-employed. That roughly 70% of all clusters enter voluntarily as they report, is broadly confirmed by the estimates of the underlying distributions in the earnings equation residual that shows a “lower tier” of under 20%. The decisions to remain in the sector are largely independent of salaried alternatives. The generalized absence of plans to expand across clusters, again does not appear related to business problems, whose incidence is similar across groups, or lack of access to credit which is rarely cited as a binding constraint. Finally, there is wide range of compliance with both social security and tax legislation: between 23% and 65% of the firms are registered with the national tax authorities, and 11% to 44% offer at least one of their workers benefits. The micro-firm sector should not be conceptually conflated with informality, although the nature of the relationship cannot be determined from this data.

ii. Results using complete variable set.

iii Factor analysis

Table 10 presents the loadings on the six factors that emerge when non-continuous variables are included. Table 11 shows that running the analysis with the residuals from the earnings equation yields extremely similar results.

Factor 1: Size/capital intensity (17.1 % of the variance). Factor 1 is essentially an augmented version of the tentative dualism factor (factor 2) of the previous analysis. In addition to loading heavily again on capital-labor ratio, earnings, hours worked and less heavily on education, it now also loads on number of workers, permanence of physical location, likelihood of paying taxes, and somewhat less reliance on credit for initial borrowing and expansion, clientele size, and being voluntarily in the sector. All of these things might be expected with more capital intensive, larger industries, who would need more permanent sites. This size and permanence, perhaps because of the need to participate more in social institutions, such as guaranteeing property rights or contracting credit, and the increased difficulty of evading may imply a rise in compliance with taxation. As in the continuous case, table 10 and figure 2 suggest a mixed distribution with the “upper tier” again, accounting for roughly 60% of the sample. But again, this does not appear to reflect dualism in the labor market: replacing earnings with η shows that, adjusted for sample wide returns to human and physical capital, earnings *decline* with positive movement along the factor.

Factor 2: Generational characteristics (10.7% of the variance). This factor is very similar to factor 1 of the previous case with the addition of a weak weightings on plans to remain in the sector. It also shares very similar underlying distributions.

Factor 3: Future plans and expansion (8.8% of the variance). This factor reveals the unsurprising fact that those who plan to expand their business are also those who plan to stay in the sector, and that these tend to have slightly higher raw earnings levels (although not adjusted), larger clients and are more voluntarily in the sector. Plans to expand are weakly correlated with problems accessing credit. Figure 2 shows that the distribution is clearly tri-modal, reflecting the discrete nature of the underlying variables and that the vast majority of the entrepreneurs are not planning to leave their present position.

Factor 4: Compliance with labor laws, financing/formality (6.9% of the variance). This factor loads most heavily on the number of workers employed and the share of workers provided the mandatory social security benefits. This could mean that larger firms less easily evade, but it could also imply that attracting more workers implies offering them benefits. The factor also loads heavily on earnings (both raw and adjusted), having used financing to start the firm, plans for additional financing, registration with tax authorities, permanence of work site, having larger clients, having no business problems, and being voluntarily in the sector.

Factor 5: Use of formal financing (5.9% of the variance). This factor loads heavily on the use of more formal sources of financing for both initiating and expanding the business. Importantly, it also loads negatively on the probability of having credit problems. Thus, those most likely to have credit problems are not those rationed out of credit markets, but precisely those who have had access before. The surprising absence of any weighting on education or experience or firm life makes it somewhat difficult to argue that either lack of information or entrepreneurial capacity prevent desire or access to financing. The factor also loads weakly on plans to expand, on larger clients, and on a larger capital-labor ratio. The distribution of the

factor is highly skewed reflecting, again, that the vast majority of firms are self sufficient in their financing, relies primarily on local clients, and has relatively low capital-labor ratios.

Factor 6 (5.7 % of the variance): Voluntary and small steady state firm size. This factor loads heavily on how voluntarily the entrepreneur entered the sector, earnings, and longevity of the firm. Associated with increased voluntariness is both lower tax payments, and less permanence in site. Factor 6 is also positively associated with client size, thus implying that those who voluntarily enter the sector may have started their firms with the express purpose of subcontracting to a specific client. The more voluntarily self-employed are more profitable--they were able to choose the project and the time. However, when including the residual in the factoring, the clientele moves negatively to the voluntary variable. This may reveal that subcontracting for a specific client does provide employment, but earnings are lower for those who respond to firm-specific demand rather than to general market demand. As interesting as the absence of any weight on financing variables, is the absence of any plans to expand or to leave the sector. Further, the factor loads negatively on experience, and positively on education. The distribution is arguably mixed although there appears to be a very small percentage of the population (14%) as involuntary, recently established, low profitability firms, with large clients. This might be consistent with individuals laid off from previous firms and now subcontracting for them.

The orthogonality of these factors suggests that many issues that are often assumed related may not be. As will be more clear in the cluster analysis, issues of credit, voluntariness, formality, and size or capital intensity are largely unrelated and cut across each other.

iib. Cluster Analysis

The extreme values of some of the discrete voluntary, financing, move-stay, and migrant variables cause clustering that effectively slices the data in a way that reveals the characteristics associated with them. For this reason, table 12 tests the direct impact of these variables on the estimated earnings equations. The mean values of, and distances between, the clusters are given in tables 13 and 14.

Clusters 2 and 4: Large, capital intensive, with and without formal financing: Clusters 2 (9.1%) and 4 (26%) are located close to one another and differ fundamentally from clusters 1,3 and 5 by having higher mean values in permanence of work site, number of employees, capital-labor ratios, and tax compliance. Cluster 2 is distinguished from all the other clusters by its reliance on formal financing for expansion and it shows the characteristics that would be predicted at the upper tail of factor 5: higher reliance on initial financing, a higher propensity to expand, more credit problems, more permanent site, greater tax compliance and the highest earnings. Table 12 suggests that these firms have a lower return to capital, as might be expected if lack of financing implied inability to build capital to where the marginal product equaled the market rate, and earn 10% more than those who do not need credit. At the other extreme, the 25% of the sample in cluster 4 represents the lower tail of factor 5: no use of formal sources of funding, either to begin or to expand, a slightly lower but still the second highest ranking in capital-labor ratio, reported earnings, level of permanence, number of workers, and size of clients, and slightly fewer credit problems. As factors 1 and 5 suggest, access to formal financing is not the critical determinant of size or mode of operation, or plans to expand.

Clusters 1,3 and 5 are similar in representing the lower tail of factor 1 in their few linkages to formal financing, small client size, fewer workers, small size, lower profitability,

lower capital-labor ratios, lack of compliance with legal norms, hours worked, and lack of a fixed physical location. But this “lower tier” is anything but homogeneous.

Cluster 1: Firm Closures (10.5%) This group differs from the entire remaining population by declaring no plans to stay self-employed, or logically, to expand and thus captures the lowest mode of the tri-modality of factor 3. The firms are distinguished by having the shortest life span, the least experience of all sectors, but a relatively high degree of voluntariness in their decision to open the business. Although their earnings are similar to those in clusters 3 and 5 who plan to remain in the sector, their high level of education suggests that these firms are kin in aspirations to the successful cluster 4, rather than to the older, less educated clusters 2 and 3. Those exiting the sector imagined that they had the potential to run a profitable business, but with experience, discover otherwise. They will leave the sector because, being young, they can.

By contrast, despite similarly low earnings, few workers, very low permanence of site, lower levels of education, the smallest clients, the lowest capital-labor ratios, and the lowest degrees of compliance with tax laws or worker benefits, both clusters 3 and 5 are committed to the sector. They present extremes along factor 6 which largely, but not exclusively, weighed on degree of voluntariness in entering the micro-firm sector.

Cluster 5: Involuntary (16%) These entrepreneurs were thrown into the sector, the least education by far, the greatest experience, but not an unusually long firm life. This suggests that they may have lost their salaried jobs at mid-life and are unable to find a new employer, a similar profile to cluster 5 in the continuous case. Their earnings adjusted for sample wide returns to human and physical capital are 24% below the rest of the sample. Returns to education and experience are lower and borderline significant, although the returns to capital are significantly higher. Thus, the most disadvantaged sector appears to be older men without the human capital

that would outweigh their advancing age, who do not have very good projects in the small firm sector or, unlike the younger workers, access to salaried jobs.

Cluster 3: Voluntary, long term, successful firms (38%). These workers entered the sector most voluntarily of all sectors. They earn less than the clusters 2 and 4 but more than workers in clusters 1 and 5 so perhaps they should be considered a third tier altogether. The fact that they have both less experience and a longer mean firm life than cluster 5 workers suggests that, perhaps faced with unappealing salaried jobs, they chose to go into business for themselves fairly early. While it would be tempting to argue that these firms are on their way to becoming firms in the top tier, their mean age is the same as those in the prosperous sectors suggesting that they are not quickly catching up. They appear only slightly interested in expanding, but report the least problems overall in operating. *What this suggests is that roughly 40% of the sector is comprised of voluntary entrepreneurs, who have reached their long run level of output at a small size and are formal neither in terms of worker protections or tax payments.*

Finally, the inclusion of the variable registering whether the entrepreneur migrated to the city as an adult was omitted because it dominated the cluster analyses. The clusters divided with one cluster high in future financing and relatively low on non-migrants, reflecting perhaps, that most who use credit to open businesses are not migrants. The remaining, four clusters represented extremes along factor 1, size and capital intensity that were, in turn, each split according to migration. These splits revealed that the characteristics associated with the 27% of the sample that migrated as adults were virtually identical to those who were native. The earnings regressions in table 12 reveal a similar story with both migrants and non-migrants showing statistically very similar earnings functions. We can conclude that migrants are not in a holding pattern for formal jobs any more than natives are.

VI. Conclusions

The continuous and complete variable sets tell a broadly similar story. Overall, the observed groupings are more suggestive of the normal heterogeneity expected in any small firm sector than the standard view of a dualistic labor market. The sector does serve as a refuge for those unable to get salaried jobs, and this group does earn somewhat less given its human capital. But it is a small minority of the population and concentrated in older workers who find it hard to be re-hired in salaried jobs if let go, rather than the young. The vast majority of entrepreneurs in all clusters is there voluntarily with desire for greater independence and higher earnings dwarfing lay-offs or inability to find a salaried job as motivation. Those leaving the sector are the youngest, but not necessarily those who entered because they could not find salaried work. They correspond more closely to the failed, but voluntary entrepreneurs that are found in any small business sector. There is little evidence that the sector serves primarily as a way of reducing labor costs for large firms through sub-contracting since only a very small minority report larger clients. Nor does not appear to be the case that migrants use the sector as a holding pattern for salaried work any more than locals do, or that they earn less given their human capital.

There is a division along the dimension of capital intensity, size, permanence, and formality, but this seems to result more from an unequal distributions of human and physical capital, factors, including perhaps, entrepreneurial talent, rather than any dualism found in the labor market: there is little relationship between capital intensity or size, and returns to human capital. Further there is little evidence that the distribution of firm size, and in particular, the small size of these firms is due to structural problems or market failures: both the optimistic, prosperous sectors and those less so show relatively similar incidence of business problems and few cite credit as the primary barrier to expansion. It is possible that the lesson has been

internalized that only with collateral and connections is access to formal sector credit remotely possible and hence most do not try. But it cannot be ruled out that the desired steady state size of over half the entrepreneurs is two or fewer workers and that they have no desire to expand further.

Table 1: Factor Decomposition
Continuous Variables Only.

	Factor 1	Factor 2	Communality
Education	-0.78	0.24	0.66
Experience	0.88	---	0.78
Firm life	0.68	0.16	0.49
Hours	0.10	0.63	0.41
K/L	---	0.67	0.46
Earnings	---	0.75	0.56
Eigenvalue	1.93	1.43	
% of variation	32.1	23.9	
<i>Moments and mixing parameter (p) of factor distributions</i>			
μ_1	-0.89	-0.36	
μ_2	0.45	1.02	
σ_1	0.56	0.69	
σ_2	0.87	1.06	
p	0.43	0.74	
LR	-159.95**	-445.80**	

**Statistically significant at the 5% level.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.57

Table 2: Factor Decomposition
Continuous Variables Only.
Earnings Residual Replacing Earnings

	Factor 1	Factor 2	Factor 3	Communality
Experience	0.88	--	--	0.78
Education	-0.79	0.22	0.10	0.68
Firm life	0.68	0.16	0.17	0.51
K/L	-0.12	0.75	0.15	0.59
Hours	--	0.74	-0.17	0.59
Residual	--	--	0.96	0.93
Eigenvalue	1.89	1.17	1.01	
% of variation	31.5	19.6	16.9	
<i>Moments and mixing parameter (p) of factor distributions</i>				
μ_1	-0.89	-0.31	-0.97	
μ_2	0.45	1.28	0.15	
σ_1	0.56	0.67	1.42	
σ_2	0.86	1.09	0.82	
p	0.34	0.80	0.13	
LR	-162.02**	-675.88**	-310.35**	

**Statistically significant at the 5% level.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.55

Figure 1: Factor Distributions – Continuous Variables Only

Table 3: Mean Cluster Centers
Continuous Variables Only.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Education	0.64	0.67	0.66	0.52	0.25	0.33
Experience	0.42	0.37	0.29	0.62	0.66	0.68
Firm life	0.22	0.12	0.10	0.66	0.15	0.45
Hours	0.53	0.53	0.42	0.54	0.45	0.46
K/L	0.12	0.61	0.06	0.56	0.07	0.05
Earnings	0.59	0.42	0.24	0.41	0.25	0.29
%	14.46	6.98	31.83	3.93	24.67	18.12
n	640	309	1409	174	1092	802

Table 4: Distance between Final Cluster Centers
Continuous Variables Only.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Cluster 1	0.00				
Cluster 2	0.54	0.00			
Cluster 3	0.41	0.60	0.00		
Cluster 4	0.69	0.61	0.86	0.00	
Cluster 5	0.58	0.77	0.55	0.78	0.00
Cluster 6	0.75	0.98	0.84	0.58	0.62

Table 5: Mean values of final cluster centers

	Cluster					
	1	2	3	4	5	6
Firm owner characteristics						
Age(years)	38.76	36.05	31.43	49.40	48.26	50.76
Experience (years)	25.07	22.03	17.54	37.16	39.32	40.82
Education (years)	7.70	8.02	7.89	6.24	2.94	3.93
Hours (worked weekly)	52.66	52.51	41.97	53.56	44.45	45.38
Firm characteristics						
Workers	2.02	1.41	1.32	1.61	1.46	1.50
Firm life (years)	7.12	4.38	3.75	7.77	5.23	22.86
Capital per worker ¹	2360.06	12,514.27	1160.00	11,441.47	1366.10	997.31
Earnings ¹	1774.86	1260.92	722.72	1253.43	773.33	868.28

¹ in 1992 pesos where 3.1 pesos = US\$1

Table 6: Motivations for Leaving Previous Job and Starting Business

Motivation	Cluster					
	1	2	3	4	5	6
<i>Why Did You Leave Your Last Job?</i>						
Laid off or business closed	13.7	26.7	19.5	9.9	19.8	13.3
Contract ended	3.3	4.6	6.3	5.7	6.4	6.0
Pay too low	25.6	18.2	31.8	17.4	25.6	25.4
To be independent	49.3	44.1	35.1	58.5	31.2	42.9
Retired	1.9	1.5	0.9	1.7	3.7	1.7
Illness or injury	1.9	1.2	2.7	0.9	7.8	4.5
Other ¹	4.1	3.8	3.7	5.9	5.4	6.2
No Response	17.7	14.9	22.0	30.3	14.5	24.7
Total	100	100	100	100	100	100
<i>Why Did You Start This Business? ²</i>						
Fired or laid off	2.8	7.4	3.5	0.6	5.1	1.5
No other work available	5.2	10.7	15.2	13.2	22.0	15.7
To be independent	52.5	41.0	41.8	59.2	42.9	44.6
Family tradition	12.7	8.4	9.9	14.4	3.6	13.3
Complement family Y	10.8	13.3	13.8	10.9	16.1	12.9
Higher pay than salaried	33.3	23.3	27.8	24.7	22.4	24.1
Flexible hours	1.7	2.3	3.7	2.9	3.5	3.5
Total Observations	640	309	1409	174	1092	802
% of sample	14.5	7.0	31.8	3.9	24.7	18.1

¹ Includes family responsibilities or moving. ² Multiple responses possible.

Table 7: Main Business Problems (in percent)¹

Problem	Cluster						Total
	1	2	3	4	5	6	
Lack of clients	53.6	48.2	63.7	56.9	62.3	61.7	60.2
Lack of credit	12.5	19.4	9.2	9.2	6.4	5.4	9.0
Lack of economic resources	18.4	22.7	21.2	13.2	23.9	17.1	20.5
Low earnings	40.9	42.7	43.6	47.7	49.0	47.8	45.4
Problems with authorities	7.2	19.7	6.7	16.1	5.6	4.4	7.4
Excessive competition	43.0	52.4	44.9	53.4	40.8	45.9	18.4
Problems with workers	2.0	2.3	0.6	2.3	0.7	1.1	1.2
Late payments	13.4	13.9	12.2	8.6	10.6	9.1	7.5
Raw materials problems	3.1	3.6	3.6	2.3	3.3	2.9	3.3

¹ Multiple responses possible.

Table8: Firm and Owner Characteristics by Cluster

	Cluster					
(as a percent of the cluster)	1	2	3	4	5	6
<i>Site</i>						
automobile	3.1	41.8	2.8	37.9	4.0	1.5
impermanent	52.3	50.8	75.7	27.0	76.6	75.8
permanent	44.5	27.8	21.5	35.1	19.5	22.7
<i>Clients</i>						
big firms	6.2	5.1	2.5	7.8	3.0	3.1
small firms	11.9	9.6	8.3	13.2	6.8	7.0
public	81.8	85.3	89.2	78.9	90.2	89.9
<i>Financing to start firm</i>						
bank	2.2	2.1	1.0	2.2	0.8	0.6
private loans	16.9	23.1	17.9	22.5	12.4	11.6
personal savings	60.9	61.5	52.5	51.1	60.4	51.4
credit (clients/suppliers)	4.7	3.8	2.3	4.8	3.4	3.7
do not need financing	15.4	9.5	26.3	19.4	23.0	32.7
<i>Additional financing</i>						
do not need	67.1	63.4	73.9	56.3	74.3	75.4
too costly	15.0	13.7	13.1	17.2	13.6	11.0
don't know how to apply	1.6	2.9	4.5	0.6	4.5	3.6
applied, did not receive	1.4	3.9	1.2	2.3	1.2	1.4
received financing from:						
bank	5.94	5.83	1.49	10.34	1.56	2.87
private loan	4.22	6.80	3.34	5.75	2.56	2.74
credit(suppliers/clients)	4.53	2.91	2.34	7.47	2.20	2.49
<i>Maquila</i>	1.7	1.0	1.5	0.6	1.6	0.9
<i>Plans</i>						
no change in firm	63.4	57.8	62.6	75.58	79.7	85.4
expand the firm	29.9	29.9	19.9	19.2	11.1	10.2
shutdown the firm	6.2	11.2	16.5	5.2	9.0	4.3
<i>Registration/tax</i>						
state treasury	31.4	30.1	18.0	36.8	19.1	18.8
Secretary of Treasury	53.8	56.0	23.2	64.9	24.7	27.6
<i>Size of last firm:</i>						
self-employed	7.7	7.8	8.3	8.6	13.7	10.2
2-5	29.1	22.0	28.7	25.9	30.9	29.6
6-10	11.7	6.5	8.2	6.9	9.2	9.4
11-15	5.3	3.9	4.2	4.6	4.4	4.1
16-50	9.2	8.4	10.7	6.3	9.2	8.0
50+	27.3	4.3	25.8	25.9	27.1	19.3
not previously employed	9.5	8.4	14.0	21.3	5.2	19.2
<i>Pay IMSS (as a % of firms with salaried workers)</i>	27.31	40.8	12.33	43.75	10.94	16.92

Table 9: Regression of log(earnings) on human and physical capital

	All	Cluster ¹					
		1	2	3	4	5	6
Education	0.03 (2.80)	0.07 (4.37)	0.05 (1.49)	0.03 (1.67)	-0.08 (-1.06)	0.05 (0.96)	-0.08** (-1.94)
Experience	0.02 (3.51)	0.01 (1.62)	0.01 (0.23)	0.03 (3.34)	0.01 (0.23)	0.06* (4.25)	-0.002 (-0.074)
Experience ² (10 ⁻⁴)	-3.07 (-5.11)	1.71** (1.66)	-0.22** (-0.079)	-7.78 (-5.03)	-2.05 (-0.47)	-8.72 (-4.86)	-1.64** (-0.67)
Ed*exp (10 ⁻⁴)	-6.95 (-2.62)	-30.00* (-5.43)	-5.30 (-0.36)	-6.45 (-0.75)	18.52 (0.31)	-17.11 (-1.49)	22.84** (2.21)
Capital (10 ⁻⁵)	1.48 (11.46)	1.15 (7.36)	0.77** (4.02)	2.03 (3.68)	1.38 (5.08)	3.1 (6.85)	4.05** (5.55)
Log(hours)	0.29 (15.84)	-0.11** (-5.43)	0.19 (3.03)	0.27 (9.48)	0.06* (0.51)	0.38** (8.95)	0.39** (8.49)
Intercept	5.08 (44.11)	7.42** (34.28)	5.68 (12.46)	5.06 (27.48)	6.53 (6.92)	4.08** (12.92)	5.41 (10.95)
dummy for cluster 1	0.86** (29.70)						
dummy for cluster 2	0.27** (6.45)						
dummy for cluster 4	0.29** (5.47)						
dummy for cluster 5	0.094** (2.84)						
dummy for cluster 6	0.24** (7.07)						
R ² adjusted	0.33	0.14	0.10	0.09	0.13	0.15	0.14
F value	201.2**	18.57**	6.79**	24.45**	5.27**	33.3**	21.9**
joint test of coefficients	10.24**	127.12**	11.74**		6.39**	3.93**	11.48**
number of observations	4426	640	309	1409	174	1092	802

Numbers in parentheses are t-statistics. ¹The asterisks on coefficients show significant differences from the coefficients of the base case, cluster 3. * = Statistically significant at the 10-percent level, ** = 5-percent level.

Table 10: Factor Decomposition
All Variables

	Factor						Communality
	1	2	3	4	5	6	
Tax	0.73	--	--	0.26	--	-0.11	0.62
Site	0.71	--	--	0.27	--	-0.12	0.60
Hours	0.62	--	--	--	--	0.10	0.40
K/L	0.58	-0.10	--	-0.14	0.24	--	0.43
Earnings	0.41	--	0.11	0.37	--	0.34	0.45
Experience	--	0.88	--	--	--	-0.16	0.80
Education	0.20	-0.77	--	--	--	0.15	0.67
Firm life	--	0.68	--	--	--	0.36	0.60
Move/stay	--	0.13	0.92	--	--	--	0.87
Expand	--	--	0.91	--	0.12	--	0.86
Workers	0.24	--	--	0.72	--	--	0.58
Benefits	--	--	--	0.72	--	--	0.53
No problems	--	--	--	0.15	-0.65	-0.12	0.47
Credit problems	--	--	-0.10	--	-0.61	0.12	0.41
Financing (new)	0.15	--	--	0.30	0.55	--	0.41
Financing (initial)	0.17	--	--	0.13	0.45	--	0.26
Voluntary	0.13	-0.13	0.14	0.18	--	0.70	0.57
Clients	0.22	--	0.11	0.27	0.11	0.48	0.38
Eigenvalue	3.07	1.93	1.58	1.23	1.05	1.02	
% of variation	17.15	10.73	8.76	6.88	5.85	5.71	
<i>Moments and mixing parameter (p) of factor distributions¹</i>							
μ_1	-0.71	-0.85	--	-0.33	--	-0.64	
μ_2	0.48	0.48	--	1.16	--	0.62	
σ_1	0.36	0.58	--	0.47	--	0.94	
σ_2	1.01	0.86	--	1.41	--	0.59	
p	0.40	0.36	--	0.78	--	0.49	
LR	-848.6**	-157.8**	--	-2225.0**	--	-308.5**	

**Statistically significant at the 5% level. Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.67

¹ Factors 3 and 6 are clearly tri-modal so moments could not be estimated.

Table 11: Factor Decomposition
All Variables
Earnings Residual Replacing Earnings

	Factor						Communality
	1	2	3	4	5	6	
Tax	0.71	--	--	0.36	--	--	0.64
Site	0.71	--	--	0.35	--	--	0.64
Hours	0.66	--	--	--	--	--	0.44
K/L	0.53	-0.10	--	--	0.26	--	0.37
Experience	--	0.87	--	0.10	--	-0.14	.080
Education	0.22	-0.77	--	--	--	0.15	0.67
Firm life	0.11	0.69	--	-0.12	--	0.28	0.58
Move/stay	--	0.13	0.91	--	--	--	0.87
Expand	--	--	0.90	0.12	0.11	--	0.86
Workers	0.21	--	--	0.69	--	0.13	0.54
Benefits	--	--	--	0.60	--	0.10	0.38
Clients	0.12	--	0.10	0.47	0.10	-0.27	0.34
No problems	--	--	--	0.20	-0.61	--	0.43
Credit problems	--	0.10	-0.12	-0.11	-0.59	0.12	0.40
Financing (new)	0.11	--	--	0.30	0.57	0.11	0.44
Financing (initial)	0.13	--	--	0.16	0.48	--	0.28
Voluntary	0.19	-0.11	--	--	--	0.71	0.56
Residual	-0.13	--	--	0.24	--	0.68	0.54
Eigenvalue	2.87	1.94	1.60	1.23	1.13	1.03	
% of variation	15.95	10.79	8.87	6.81	6.27	5.75	
<i>Moments and mixing parameter (p) of factor distributions¹</i>							
μ_1	-0.70	-0.85	--	-0.36	--	-0.54	
μ_2	0.46	0.49	--	1.09	--	0.54	
σ_1	0.37	0.58	--	0.58	--	1.05	
σ_2	1.02	0.86	--	1.19	--	0.56	
p	0.40	0.37		0.75		0.50	
LR	-767.9**	160.3**		-1124.3**		-415.9**	

**Statistically significant at the 5% level. Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.65

¹Factors 3 and 6 are clearly tri-modal so moments could not be estimated.

Figure 2: Factor Distributions – All Variables

Table 12: Regression of log(earnings) on human/physical capital and possible segmentation variables

	need more credit		stay/leave		migrant status		voluntariness	
	<i>i</i>	<i>ii</i>	<i>iii</i>	<i>iv</i>	<i>v</i>	<i>vi</i>	<i>vii</i>	<i>viii</i>
Education	0.04 (4.29)	0.04 (3.76)	0.04 (4.27)	0.04 (3.70)	0.04 (4.39)	0.04 (3.20)	0.05 (4.74)	0.05 (5.05)
Experience	0.03 (6.1)	0.03 (5.87)	0.03 (5.92)	0.03 (5.90)	0.03 (6.17)	0.06 (4.20)	0.03 (6.55)	0.04 (6.61)
Exp ² (10 ⁻⁴)	-5.14 (-7.82)	-5.35 (-7.69)	-5.06 (-7.73)	-5.52 (-7.94)	-5.19 (-7.89)	-4.29 (-5.49)	-5.23 (-8.02)	-5.45 (-7.72)
Ed*Exp (10 ⁻⁴)	-6.20 (-2.13)	-6.25 (-2.02)	-5.77 (-1.99)	-4.69 (-1.54)	-0.06 (-2.16)	-3.16 (-0.89)	-7.58 (-2.62)	-0.001 (-3.23)
Capital (10 ⁻⁵)	1.80 (15.23)	1.98 (13.84)	1.86 (16.13)	1.83 (15.53)	1.88 (16.26)	1.86 (13.28)	1.83 (15.91)	1.73 (14.55)
Log(hours)	0.39 (19.47)	0.41 (18.97)	0.38 (19.02)	0.36 (16.84)	0.39 (19.49)	0.40 (16.88)	0.38 (19.29)	0.38 (17.35)
Intercept	4.58 (36.49)	4.53 (34.16)	4.66 (37.14)	4.72 (34.82)	4.57 (36.38)	4.62 (31.97)	4.59 (36.91)	4.55 (33.91)
Dummy ¹	0.10 (3.42)	0.29 (0.68)	-0.21 (-7.11)	-0.58 (-1.63)	0.001 (0.07)	-0.34 (-1.09)	-0.24 (-9.16)	0.22 (0.61)
Dummy interactive								
Education		0.04 (1.33)		0.03 (0.89)		0.03 (1.13)		-0.05 (-1.73)
Experience		0.004 (0.21)		-0.02 (-1.02)		0.03 (2.16)		-0.02 (-1.61)
Exp ² (10 ⁻⁴)		1.27 (0.58)		4.50 (2.09)		-3.63 (-2.27)		2.35 (1.27)
Ed*Exp (10 ⁻⁴)		-5.85 (-0.61)		-9.58 (-0.96)		-10.05 (-1.50)		16.56 (2.00)
Capital (10 ⁻⁵)		-0.58 (-2.29)		0.03 (0.62)		0.06 (0.26)		1.61 (3.46)
Log(hours)		-0.15 (-0.25)		0.11 (1.84)		-0.04 (-0.87)		0.01 (0.14)
R ² adjusted	0.19	0.19	0.19	0.20	0.19	0.19	0.20	0.20
F value ²	11.70	4.50	50.59	10.19	0.016	0.93	83.88	14.60

Numbers in parentheses are t-statistics. Number of observations = 4426 for each regression.¹ The dummy corresponds to the segmentation variable being tested in each column. It takes a value of 1 if credit is a constraint (regressions *i*, *ii*), if the firm owner wants to shutdown his business (regressions *iii*, *iv*), if the firm owner migrated as an adult (regressions *v*, *vi*), and if the firm owner involuntarily entered the small firm sector (regressions *vii*, *viii*), 0 otherwise.² where the degrees of freedom for the regressions with only a dummy term (*i*, *iii*, *v*, *vi*) are (1, 4418) and the degrees of freedom for the regressions with all the dummy interactive terms (*ii*, *iv*, *vi*, *viii*) are (7, 4412).

Table 13: Cluster Centers
All Variables

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Education	0.54	0.55	0.47	0.56	0.36
Experience	0.41	0.50	0.47	0.49	0.59
Firm life	0.16	0.30	0.30	0.27	0.27
Voluntary	0.64	0.87	0.98	0.89	0.18
Hours	0.42	0.52	0.43	0.54	0.41
K/L	0.11	0.24	0.08	0.19	0.08
Earnings	0.25	0.44	0.31	0.38	0.23
Site	0.25	0.68	0.10	0.85	0.16
Workers	0.10	0.37	0.09	0.26	0.07
Benefits	0.03	0.15	0.02	0.08	0.02
Tax	0.16	0.60	0.05	0.63	0.09
Move/stay	0.00	0.95	1.00	0.99	1.00
Expand	0.00	0.68	0.58	0.64	0.56
Clients	0.21	0.33	0.19	0.29	0.20
No problems	0.11	0.09	0.20	0.19	0.18
Credit problems	0.94	0.81	0.96	0.92	0.96
Financing (initial)	0.08	0.24	0.05	0.11	0.07
Financing (new)	0.06	1.00	0.01	0.00	0.06
%	10.48	9.08	38.05	25.62	16.76
n	464	402	1684	1134	742

Table 14: Distance between Final Cluster Centers
All Variables

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	0.00			
Cluster 2	1.71	0.00		
Cluster 3	1.24	1.37	0.00	
Cluster 4	1.46	1.05	1.00	0.00
Cluster 5	1.27	1.49	0.82	1.20

Appendix: Construction of Variables

A. Personal Characteristics:

1. *Education*. The responses were categorized as count variables where 0 is no education, 1 is one year of education, and so forth up to 12 is completed high school. Those with more than a high school diploma were omitted since we are primarily concerned with the relatively less educated entrepreneurs.

2. *Work Experience of the Entrepreneur*. (Experience) This is total work experience in any sector and is calculated as age - 6 - years of education.

3. *Why initiated enterprise*. (Voluntary) This variable is intended to capture the how voluntarily the entrepreneur entered the sector by using the motivational responses. The variable is coded into three categories: involuntary (to supplement family income, couldn't find other job, fired from last job) coded at 0, semi-voluntary (family tradition), coded at 0.5, and voluntary (to be independent, to earn a higher salary than possible if employed, flexible working hours) coded at 1.

B. Characteristics of the firm:

1. *Firm life*. Measured in years. The data showed a vast number of very young enterprises and then an extremely long tail encompassing successful long-lived firms. Simply normalizing the data compressed 90% of the observations into the range [0,0.39]. The apparent lack of variation diminishes the power of this variable to drive clusters. Thus, the values were recoded into deciles in increments of 2.5 years with an upper bound of 1 representing all individuals with firm lives greater than 25 years.

2. *Capital-Labor Ratio*. (K/L) This variable divides the reported total capital stock of the firm by the total number of workers in the firm. This variable also showed some extreme outliers so the top 5% were omitted.

3. *Earnings* Profits net of all expenses. The existence of extreme outliers meant that on normalization, 95% of the sample is in the range [0,0.04] leaving little variance for the bulk of the sample around which to factor or cluster. By omitting the top 5%, approximately 95% of the new sample is in the range [0,0.71].

4. *Permanence of Worksite* (Site) takes a value of 1 if the firm has a permanent locale, 0.5 if the firm can be moved, but it has a specific work site as recognized by the other firm owners and its customers and 0 if the firm does not have a permanent work site.

5. *Hours Worked* (Hours) is average number of hours worked per week.

6. *Total number of workers* (Workers) counts the number of workers (including the owner) working in that firm. Since 68.1% of the firms are self-employment (1 worker) and 86.7% of the firms have only one employee while 0.1% of the firms have eight employees, the

top end of the distribution is collapsed into the value for the upper bound, thus stretching left side of the distribution.

C. Entrepreneurial Dynamics:

1. *Permanence in the Sector* (Move/stay) takes a value of 1 if the entrepreneur reports that he plans to remain in the sector, and 0 if he plans to look for salaried work.

2. *Plans to Expand* (Expand) takes a value of 0 if the Move/stay variable takes a value of 0, 1 if the entrepreneur plans to expand the line of products, undertake new investments or hire more workers, and 0.5 if he plans to stay in business, but not expand.

3. *Clients* captures the degree of personalism clientele. The responses were categorized into three groups: large formal clients (large businesses, large factories) coded as 1, smaller less formal clients (small businesses, small factories or workshops) coded as 0.66, informal clients (the public) coded as 0.33 and family are coded as 0.

4. *Initial financing for firm* (Financing-initial) captures where the funds to start the business were obtained. There were nine possible responses in the survey that were reclassified into three general categories: formal credit markets (banks, or savings banks) coded as 1, semi-formal credit markets (credit from clients, credit from suppliers) coded as 0.5, and informal credit markets (loans from friends or relatives, private loans, savings, severance payments) coded as 0.

5. *Demand for additional financing* (Financing-new) captures whether the firm has sought additional financing from any source.

6. *Financing problems*. (Credit Problems) codes how problematic credit is. It takes a value of 1 if credit is not cited as a business problem, 0.66 if credit is the third most pressing problem, 0.33 if it is the second most important problem, and 0 if credit is the primary problem in the firm.

7. *No Problems* takes a value of 1 if the entrepreneur states that his firm is free of problems and a 0 if the owner does not face market, institutional, or resource problems.

D. Participation in formal market and legal institutions

1. *Registration with Tax Authorities* (tax) This variable could take a value of 0 if not registered at all, 0.5 if registered with either the state or municipal authorities or with the federal treasury, and 1 if both. Though often used as the critical definition of formality, we treat it as only one of several firm characteristics.

2. *Compliance with Labor Legislation*. (Benefits) By law, all salaried workers are to be registered in the national health benefits program. The fraction of covered workers is a measure of one dimension of "formality." However, because the vast majority of the sample are self-employed and have no employees (65%) and entrepreneurs are not required to pay IMSS for themselves or their unpaid workers, this variable is relevant to only a fraction of the observations.

VII. References

- Anderberg, M.R. (1973), *Cluster Analysis for Applications*, Academic Press, NY.
- Anderson, K. H., J. S. Butler and F. A. Sloan (1987), "Labor Market Segmentation: A Cluster Analysis of Job Grouping and Barriers to Entry." *Southern Economic Journal*, 53:3 January.
- Aronson, R.L.(1991), *Self-Employment, A Labor Market Perspective*. ILR Press, Ithaca.
- Balán, J. and H.L. Browning, E. Jelin (1973), *Men in a Developing Society*, Institute of Latin American Studies, U.T. Austin.
- Beneria, L.(1989) Subcontracting and Employment Dynamics in Mexico City, in Portes et al., *The Informal Economy*, Johns Hopkins.
- Castells, M. and A. Portes (1989), "World Underneath: The Origins, Dynamics, and Effects of the Informal Economy," in A., M. Castells, and L. A. Benton, eds., *The Informal Economy: Studies in Advanced and Less Developed Countries*, Baltimore, Maryland: Johns Hopkins University Press.
- Davis, S. J., J. Haltiwanger, and S. Schuh (1994), "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts," in L. C. Solmon and A. R. Levenson, eds., *Labor Markets, Employment Policy, and Job Creation*, Boulder, Colorado: Westview Press.
- de Soto, H. (1989), *The Other Path: The Invisible Revolution in the Third World*, (New York: Harper and Row).
- Dunne, T., M. J. Roberts, and L. Samuelson (1989), "The Growth and Failure of U.S. Manufacturing Plants," *Quarterly Journal of Economics*, 104:4, 671-698.
- Elizondo, N. (1992), "Illegality in the Urban Informal Sector of Mexico City," in V.E. Tokman, ed., *Beyond Regulation, The Informal Economy in Latin America*, Lynne Rienner, Boulder, Colorado.
- Esfahani, H. and D. Salehi-Isfahani (1989), "Effort Observability and Worker Productivity: Towards and Explanation of Economic Dualism," *Economic Journal*, 99, 818-836.
- Evans, D. S. and L. S. Leighton (1989), "Some Empirical Aspects of Entrepreneurship," *American Economic Review*, 79:3, 519-535.
- Everitt, B. (1974) *Cluster Analysis*, John Wiley and Sons, NY.
- Fairlie, R. W. and B.D. Meyer, "Trends in Self-Employment Among White and Black Men: 1910-1990," Mimeo, U.C. Santa Cruz.

Fields G. S. (1990), "Labor Market Modeling and the Urban Informal Sector: Theory and Evidence," in OECD, *The Informal Sector Revisited*, Paris.

Flatau, P. R. and Philip E. T. Lewis, Segmented Labour Markets in Australia, *Applied Economics*, 25, 285-294.

Gregory, P. (1986), *The Myth of Market Failure: Employment and the Labor Market in Mexico*, Baltimore: World Bank and Johns Hopkins.

Harman, H.H. (1976), *Modern Factor Analysis*, 3rd ed, University of Chicago Press, Chicago.

Harris, J.R. and M. P. Todaro (1970), "Migration, Unemployment, and Development: A Two Sector Analysis," *American Economic Review*, 60:1, 126-142.

Hart, K. (1972), *Employment, Income and Inequality: A Strategy for Increasing Productive Employment in Kenya*, Geneva, ILO.

Jovanovic, B. (1982), "Selection and Evolution of Industry," *Econometrica*, 649-670.

Kim, J. and C. W. Mueller (1978), *Statistical Methods and Practical Issues*, Sage Publication, Beverly Hills.

Levenson, A.R. and W. F. Maloney (1997), "The Informal Sector, Firm Dynamics and Institutional Participation" Mimeo, University of Illinois and the Milken Institute.

Loayza, N.V. (1994) "Labor Regulations and the Informal Economy," Policy Research Working Paper 1335, World Bank, August.

Lopez-Castano, H. (1990) "Inestabilidad Laboral y Ciclo de Vida en Colombia," (Labor Instability and the Life-Cycle in Colombia) *Coyuntura Economica*, 20:1, p 173-191.

Lucas, R. E., Jr. (1978), "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9:2, 508-23.

Maloney, W. (1996) "Are LDC Labor Markets Dualistic?" University of Illinois Working Paper.

Maloney W. (1997) "Labor Market Structure in Mexico: Time Series Evidence on Competing Views" University of Illinois Working Paper

Marcouiller, D. V., Ruiz de Castilla and C. Woodruff (forthcoming), " Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador, and Peru", *Economic Development and Cultural Change*.

Marquez, C. and J. Ros (1990), "Segmentacion del Mercado de Trabajo y Desarrollo Economico en Mexico," (Labor Market Segmentation and Economic Development in Mexico) *El Trimestre Economico*, Abril-Junio.

Mazumdar, D. (1981), *The Urban Labor Market and Income Distribution*, Oxford University Press, New York.

Mazumdar, D. (1983). "Segmented Labor Markets in LDC's." *American Economic Review*, 73, 2254-259.

Oster, G. (1979), A Factor Analysis Test of the Theory of the Dual Economy, *Review of Economics and Statistics*, 33-41.

Ozorio de Almeida, A.L., and S.E.M Graham and L. F. Alves (1994) "Poverty, Deregulation and Informal Employment in Mexico," ESP Working Paper, World Bank, July.

Peattie, L. (1982), "What is to be Done with the 'Informal Sector'? A Case Study of Shoe Manufacturers in Colombia" in H. Safa ed, *Towards a Political Economy of Urbanization in Third World Countries*, Delhi:Oxford University.

Peattie, L. (1987) "An Idea in Good Currency and How It Grew: The Informal Sector," *World Development*, 15:7 851-860.

Portes, A., M. Castells, and L. A. Benton, eds. (1989), *The Informal Economy: Studies in Advanced and Less Developed Countries*, (Baltimore: Johns Hopkins University Press).

Portes, R. and R. Schauffler (1993), "Competing Perspective on the Latin American Informal Sector," *Population and Development Review*, 19:1, 33-59.

Rauch, J. E. (1991), "Modeling the Informal Sector Formally," *Journal of Development Economics*, 35:1, 33-47.

Roberts, B. (1989) "Employment Structure, Life Cycle and Life Chances: Formal and Informal Sectors in Guadalajara," in Portes et al., *The Informal Economy*, Johns Hopkins.

Secretaria del Trabajo y Prevision Social (1993), *El Sector Informal en Mexico* (The Informal Sector in Mexico), no 2.

Sloane, P. J., P. D. Murphy, I. Theodossiou and M. White (1993), "Labour Market Segmentation: a Local Labour Market Analysis using Alternative Approaches." *Applied Economics*, 25, 569-581.

Späth, H. (1980), *Cluster Analysis Algorithms for Data Reduction and Classification of Objects*, John Wiley and Sons, NY.

Stiglitz, J. E., "Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model," *Quarterly Journal of Economics*, 88:194-227.

Tokman, V.E. (1989), "Policies for a Heterogeneous Informal Sector in Latin America," *World Development*, July.

Tokman, V.E. (1992) *Beyond Regulation, The Informal Economy in Latin America*, Lynne Rienner, Boulder.

Turnham D. and D. Eröcal (1990), "Unemployment in Developing Countries, New Light on an Old Problem," Technical Paper, OECD Development Centre.

U.S. Dept of Labor (1992) *The Informal Sector in Mexico*.

U.S. Dept of Labor (1993) *Work Without Protections: Case Studies of the Informal Sector in Developing Countries*.

World Bank (1995), "Labor and Economic Reforms in Latin America and the Caribbean," *Regional Perspectives on World Development Report*.

Policy Research Working Paper Series

Title	Author	Date	Contact for paper
WPS1979 Banking on Crises: Expensive Lessons from Recent Financial Crises	Gerard Caprio, Jr.	September 1998	P. Sintim-Aboagye 38526
WPS1980 The Effect of Household Wealth on Educational Attainment: Demographic and Health Survey Evidence	Deon Filmer Lant Pritchett	September 1998	S. Fallon 38009
WPS1981 Evaluating Public Expenditures When Governments Must Rely on Distortionary Taxation	James E. Anderson Will Martin	September 1998	L. Tabada 36896
WPS1982 Analyzing Financial Sectors in Transition: With Special Reference to the Former Soviet Union	Alan Roe Paul Siegelbaum Tim King	September 1998	D. Cortijo 84005
WPS1983 Pension Reform in Small Developing Countries	Thomas Charles Glaessner Salvador Valdés-Prieto	September 1998	M. Navarro 84722
WPS1984 NAFTA, Capital Mobility, and Mexico's Financial System	Thomas Charles Glaessner Daniel Oks	September 1998	M. Navarro 84722
WPS1985 The Optimality of Being Efficient: Designing Auctions	Lawrence M. Ausubel Peter Cramton	September 1998	S. Vivas 82809
WPS1986 Putting Auction Theory to Work: The Simultaneous Ascending Auction	Paul Milgrom	September 1998	S. Vivas 82809
WPS1987 Political Economy and Political Risks of Institutional Reform in the Water Sector	Ariel Dinar Trichur K. Balakrishnan Joseph Wambia	September 1998	F. Toppin 30450
WPS1988 The Informal Sector, Firm Dynamics, and Institutional Participation	Alec R. Levenson William F. Maloney	September 1998	T. Gomez 32127
WPS1989 Contingent Government Liabilities: A Hidden Risk for Fiscal Stability	Hana Polackova	October 1998	A. Panton 85433
WPS1990 The East Asia Crisis and Corporate Finances: The Untold Micro Story	Michael Pomerleano	October 1998	N. Dacanay 34068
WPS1991 Reducing Air Pollution from Urban Passenger Transport: A Framework for Policy Analysis	Mark Heil Sheoli Pargal	October 1998	R. Yazigi 37176
WPS1992 The Present Outlook for Trade Negotiations in the World Trade Organization	John Croome	October 1998	L. Tabada 36896

Policy Research Working Paper Series

Title	Author	Date	Contact for paper
WPS1993 Financial Safety Nets and Incentive Structures in Latin America	Philip L. Brock	October 1998	K. Labrie 38256
WPS1994 Estimating Wealth Effects without Expenditure Data — or Tears: with an Application to Educational Enrollments in States of India	Deon Filmer Lant Pritchett	October 1998	S. Fallon 38009
WPS1995 What Macroeconomic Policies Are "Sound?"	Mansoor Dailami Nadeem ul Haque	October 1998	B. Nedrow 31585
WPS1996 Namibia's Social Safety Net: Issues and Options for Reform	Kalinidhi Subbarao	October 1998	P. Lizarondo 87199
WPS1997 On Measuring Literacy	Kaushik Basu James E. Foster	October 1998	M. Mason 30809
WPS1998 The Structure and Determinants of Inequality and Poverty Reduction in Ghana, 1988–92	Sudharshan Canagarajah Dipak Mazumdar Xiao Ye	October 1998	A. Garscadden 38400