Labor Earnings Dynamics with a Large Informal Sector

Diego B. P. Gomes*

Felipe S. Iachan[†]

Cezar Santos[‡]

February 14, 2019

Preliminary.

Abstract

We study labor earnings dynamics in a developing economy with a large informal sector. We use nationally representative Brazilian panel data that cover both formal and informal workers. We provide two main contributions. First, we document large disparities in earnings fluctuations faced across these segments of the labor market, as well as the high intensity of transitions between them. Informality is associated with more volatile earnings, while agents in the formal sector are subject to significant downside risk. Transitions across formal and informal employment bring large earnings shocks on average and have a frequency that depends on age and the initial earnings level. Second, we assess the consequences of these empirical disparities on decisions of consumption and savings. Our tool is a standard life cycle model with heterogeneous agents where the earnings processes are estimated to reflect the aforementioned empirical findings. The simulations reveal sizable impacts.

JEL codes: E24, J24, J31

Keywords: Earnings dynamics, higher-order earnings risk, non-Gaussian shocks, informality

*University of Alberta, diego.gomes@gmail.com

[†]FGV EPGE – Escola Brasileira de Economia e Finanças; felipe.iachan@fgv.br.

[‡]FGV EPGE – Escola Brasileira de Economia e Finanças and Banco de Portugal; cezarsantos.econ@gmail.com

1 Introduction

Uninsurable earnings risk influences many individual decisions and takes center stage in contemporary Macroeconomics. In the volatile economies of developing countries, it is reasonable to expect these risks to be both larger and more impactful. While a good match with an expanding firm can translate into accelerated experience accumulation and wage gains, a negative shock, as a layoff from a waning economic activity, can lead to loss of formal employment and a difficulty in ever recovering the same level of income.

Recent research with large panel datasets has emphasized the importance of accounting for a rich set of features of the empirical behavior of labor earnings risk.¹ Income risk is pervasive, but the availability of appropriate data has been a constraint and most of what we have learned refers to a select group of developed economies. In this paper, we first document key facts about labor earnings dynamics in a developing economy. We use Brazilian panel data from the first quarter of 2012 to the second quarter of 2018, that cover the whole country. As in countries with similar income levels, informality plays a large role in the labor market and a failure to account for it leads to a distorted description of labor dynamics.²

An informal worker is defined as an employee that lacks the required registration into social security systems.³ As such, there is no compliance with statutory labor rights and obligations. For example, there are no patronal contributions to social security, no layoff compensation, no access to unemployment insurance upon job separation, and, potentially, no paid vacations. While on the one hand, conditions of informal workers are generally considered worse than in the formal sector, on the other hand, temporary informal employment can serve as an insurance mechanism.

We use non-parametric methods to account for heterogeneity in labor earnings risk across age groups, current income level, and status of employment and to describe the likelihood of transitions between these. We find many striking differences between the nature of earnings risk faced by workers under formal and informal labor arrangements.

First, as expected, formal sector workers typically enjoy both higher levels of earnings and less

¹See, for example, Arellano, Blundell, and Bonhomme (2017), De Nardi, Fella, and Pardo (2016), and Guvenen, Karahan, Ozkan, and Song (2016).

²International Labor Organization (2013) compiled statistics on employment in the informal economy from 47 medium and low-income countries. In 15 countries, informal employment represents at least two thirds of total non-agricultural employment, and in 24 countries it exceeds 50 per cent.

³In particular, a formal worker in Brazil has a required registration in *Carteira de Trabalho* (a passport-like worker ID) and *Relação Anual de Informações Sociais* (RAIS), a required filing with the Labor Ministry.

volatile innovations. The standard deviation of log-earnings innovations⁴ is 0.47 log-points for agents that are initially in the formal sector, while it is 0.70 for agents that are in the informal sector. However, the nature of the risk they face is markedly different. For instance, a striking observation at first glance is that workers that are in the formal sector at a given moment in time are subject to, on average, earnings losses over a one-year period (-0.04 log-point innovations). Informal workers have, on average, zero gains. Behind these puzzling averages, lie great asymmetries. For instance, formal workers are subject to significant downside risk, which is embedded in transitions to the informal sector.

A transition of this form is a significant negative shock: on average, a worker that switches to informality across two periods loses 0.28 log-points of net earnings. These innovations are reasonably heterogenous (a conditional standard deviation of 0.70) and also significantly left-skewed (-0.74), emphasizing the role of further downside risk even within the group that suffers the negative shock. On the other direction, a switch from the informal sector toward formal employment is somewhat a mirror reflection of these shocks. A worker typically gains 0.18 log-points, with a distribution that has a conditional standard deviation of (0.67) and is positively skewed (0.51).

We also document that earnings innovations show large deviations from normality and study how the distribution of these shocks depends on individual characteristics, such as the level of income and age. Our main finding is that transitions across sectors are very frequent (FIXME: can we aggregate for the intro?) and age dependent. Young agents are much more likely to transition across sectors. Also, it is typically the low earning workers of the formal sectors that fall to informality more easily, which it is the high-earnings workers from the informal sector that most frequently switch to formal employment.

These rich patterns and the heterogeneity in exposure to labor income risks are expected to influence consumption and savings decisions. Unfortunately, there is no longitudinal data available to directly study these decisions. Thus, we apply a quantitative model to predict their optimal responses to labor income risk. As an input to the model, we use the empirical patterns of earnings risk and transitions across sectors. We also make these summarized data publicly available for future use.

From the structural model, we find substantial differences across formal and informal workers. Formal workers tend to consume more due to their higher income and their consumption grows faster over the life-cycle. Interestingly, the wealth to income ratio is higher for informal workers. They choose higher wealth because they face a riskier earnings process. Finally, we also compare our baseline calibration with

⁴We study innovations in residuals of log-net-earnings, after deflating, imputing taxes and mandatory social security contributions due, and removing age and year effects.

what would happen if formal workers never faced the possibility of transitioning to the informal sector. This is essentially what a researcher would find if she looked at formal workers only. Formal workers without transition risk to the informal sector would save substantially less since their precautionary reason to save is not as strong.

This paper connects to four strands of the literature. First, there are studies documenting labor earnings risk in a variety of countries. Meghir and Pistaferri (2011) provide a survey of the literature studying the U.S., the country with the absolute majority of articles. Other developed countries have also been studied, like the U.K. (Dickens, 2000), Canada (Baker and Solon, 2003), Italy (Cappellari, 2004), Spain (Alvarez, 2004), and Germany (Krebs and Yao, 2016). These works usually study developed countries and do not focus on the differences between formal and informal sectors.

A second set of recent papers use rich micro data sets to study nonlinear and higher-order aspects of labor earnings risk, such as Arellano et al. (2017), De Nardi et al. (2016), and Guvenen et al. (2016). We also document high-order risk, but add to this literature by focusing on the dynamics of labor earnings of the formal and informal sectors separately.

Processes for labor earnings risk is a natural input into quantitative macroeconomic models. See for instance, canonical applications for consumption (Krueger and Perri, 2006), savings (Aiyagari, 1994), human capital accumulation (Krebs, 2003), and labor supply (Heathcote, Storesletten, and Violante, 2010). These risks also relate to macroeconomic variables such as the risk-free rate (Huggett, 1993) and asset prices (Constantinides and Duffie, 1996). We contribute to this literature by studying the consumption and savings decisions of formal and informal workers using a workhorse quantitative life-cycle model.

Finally, there is also a literature that focuses on informality per se. Meghir, Narita, and Robin (2015) study informal firms in Brazil. Ulyssea (2017) allows informal firms to coexist with formal firms that may hire informal workers. Both papers have homogeneous workers that are not subject to income shocks. Other papers focus on the relationship between informality and trade (Coşar, Guner, and Tybout (2016)), tax collection (Ordóñez (2014)), job creation and destruction (Bosch and Esteban-Pretel (2012)), economic development (Porta and Shleifer (2008)), managerial ability (De Paula and Scheinkman (2011)), and regulation (Rocha, Ulyssea, and Rachter (2018)). None of these papers focus on differences in labor earnings risk across formal and informal sectors like we do.

This paper is organized in four additional sections besides this introduction. Section 2 describes the data. Section 3 discusses the empirical findings, whereas Section 4 reports the results from the

Table 1: **Sample description**

	Men		Women		Total	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
F-F	166,415	43.4	129,921	51.3	296,336	46.5
F-I	29,384	7.7	19,443	7.7	48,827	7.7
I-F	28,906	7.5	21,820	8.6	50,726	8.0
I-I	158,719	41.4	82,170	32.4	240,889	37.8
Total	383,424	100	253,354	100	636,778	100

This table presents the number of occurrences for all possible combinations of forms of employment in the initial and terminal observations of an individual. F denotes formal employment, while I denotes informal employment.

structural model. Finally, Section 5 concludes.

2 Data

We use Brazilian survey data from the *Pesquisa Nacional por Amostra de Domicílios Contínua* (PNADC) ranging from 2012q1 to 2018q2. This survey is conducted by the *Instituto Brasileiro de Geografia e Estatística* (IBGE), which is the agency responsible for official collection of statistical information in Brazil. The PNADC aims to produce indicators that monitor quarterly fluctuations in the workforce. It has national coverage and considers both the formal and informal sectors.

In each quarter, PNADC surveys around 211,000 households in approximately 16,000 census tracts. It follows a rotation scheme known as 1–2(5), where each household is visited five times during five consecutive quarters. The second visit occurs three months after the first one, the third visit happens three months after the second, and so on. From one quarter to the next, there is an overlap of 80% of households, and between the same quarter in two consecutive years, an overlap of 20%. This last feature allows us to construct a panel dataset that keeps track of individuals in an interval of one year.

Our final sample considers workers aged between 18 and 65, focusing on an age range where individuals have a strong labor market attachment. We kept only employed individuals with positive earnings in two consecutive years. The raw earnings measure is the monthly gross income from the main job. Real earnings were calculated in 2018q3 *Reais* (R\$) using the Brazilian price index *Índice Nacional de Preços ao Consumidor Amplo* (IPCA). We used the monthly and regional indexes provided by IBGE, taking into account regional differences in the inflation rate.

To assess if a worker belongs to the formal or informal sector, we use information about an employment

Figure 1: Sample size over time

Notes: The values on the horizontal axis refer to the first quarter of a one-year period. For example, the values under “2016q1” refer to the number of workers present in both 2016q1 and 2017q1.

entry in their *Carteira de Trabalho e Previdência Social* (CTPS). This document is issued by the Brazilian Ministry of Labor and is required for all formal employment relationships in the private sector. It contains information on the characteristics of the job, such as compensation, and must be signed by both the employer and the employee. Thus, our classification of formal workers involves the employees with a CTPS, plus employees in the public sector and armed forces.. Complementarily, informal workers are the employees without CTPS, those self-employed, and the auxiliary family workers. We exclude the small mass of employers from the analysis.

Table 1 presents the number of observations in the pooled sample. The total number of workers in two consecutive years amounts to more than half million, of which around 60% are men. Figure 1 breaks down the sample over time and the form of employment. Transitions across formal and informal employment are stable and non-trivial. These transitions are the main focus of the analysis in Section 3.2.

3 The earnings dynamics of formal and informal workers

We first study the distribution of the level of earnings across the two segments of the labor market. As Figure 2a illustrates, workers under a formal employment arrangement fare significantly better. Their conditional distribution of earnings dominates, in a first-order stochastic sense, the earnings distribution

Figure 2: The behavior of net earnings, conditional of form of employment

Notes: The panel on the left displays the inverse cumulative distribution of net earnings: on the horizontal axis we have shares between 0 and 100%, while the curves display the respective quantiles of the distribution of earnings (conditional on the sector of employment). The panel on the right displays the densities of the earnings distributions for the two sectors of employment and moments associated with these distributions.

of informal workers. This difference is expected given previous studies of the Brazilian labor market.⁵ Much less is known about the risky evolution of individual labor earnings in developing economies. What are the key features of the growth trends and embedded risk in the earnings of workers? How do transitions across sectors affect their earnings perspectives? Which agents are specially subject to these transitions?

We focus on disposable yearly income (net earnings), as common in the literature. When studying formal workers, we subtract imputed income taxes and mandatory social contributions, based on contemporaneous tax rates and brackets, and account for a deferred 13th monthly salary mandated by Brazilian labor laws.⁶ When studying informal sector workers, we subtract imputed social security contributions from the interviewees that report a contribution.⁷ We additionally remove year and age fixed effects. The procedure is detailed in Appendix A. The main lessons are largely unchanged when studying gross income.

Following the standard in the literature, we adopt the *log-earnings change* measure.⁸ Let $y_{i,t}$ be the (residual net) earnings of worker i in period t . Then, the log-earnings change over a period of one year is defined as

$$\Delta \ln y_{i,t} \equiv \ln y_{i,t+1} - \ln y_{i,t}.$$

⁵See, for example, Botelho and Ponczek (2011).

⁶More details are available in Law 13.467/2017.

⁷Informal sector workers can voluntarily contribute in the *autonomous worker* category.

⁸See, for instance, Guvenen, Ozkan, and Song, 2014, and Guvenen et al., 2016.

We calculate it for all workers in all one year periods of our pooled sample. From now on, we concentrate on the results for men, for comparability to earlier work. When we restrict attention to women participating in the labor force, results are similar.

Figure 2b plots the probability density of log-earnings changes, conditioning on sector of participation at date t . A few features are noteworthy. First, agents in the formal sector seem to be subject to significant downside risk in their income processes, as evidenced by a negative mean change (-0.04 log points) and negative skewness (-1.08). This downside risk is largely explained by the presence of transitions out of formal employment, as we document further in the next sections.

At the same time, the earnings of agents originally in the informal sector are more volatile, as indicated by the standard deviation 0.7, compared with a lower 0.47 for agents originally in the formal sector. Unlike for formal sector workers, the risk faced by informal laborers is symmetric (its skewness is roughly 0). There is also a noticeable difference in kurtosis (15.71 for formal sector workers, 6.60 for informal sector workers, both exceeding the normal distribution), which is the subject of further study in the next sections. There, to dig deeper into the nature of income risk faced by individuals in the two labor market segments, we separate sector stayers and agents experiencing a sector transition.

3.1 Sector Stayers

We now focus on stayers, workers who are observed in the same sector of employment over two one-year-apart surveys. Notice, for instance, that formal sector stayers is the single group typically covered by high quality administrative data.⁹ One of our main objectives is to document how a labor income dynamics description based solely on data covering the formal sector would miss important features of the earnings risk faced by workers in a developing economy.

Figure 3 explores the shape of the density functions for sectors stayers. We can observe large deviations from normal distributions with the same mean and variance, with a typical case of high kurtosis: a central spike, thin shoulders around this spike, and heavy tails. This feature is especially pronounced for formal sector stayers, whose kurtosis for log-earnings changes is 21.06.¹⁰ Informal sector stayers, in their turn, are associated with a kurtosis is 6.42. Although shadowed by the behavior of the formal sector's counterpart, this is still significantly higher than the one displayed by the normal distribution. Both distributions are slightly left skewed (skewness of -0.71 and -0.11, respectively),

⁹In the particular case of Brazil, the RAIS database from the Ministry of Labor.

¹⁰The value of the kurtosis of a normal variable is 3.

emphasizing the importance of a thick tail of negative shocks. Guvenen et al. (2016) and De Nardi et al. (2016) also find large deviations from normality in the U.S. labor market, with the former using administrative Social Security data and the latter using the Panel Study of Income Dynamics (PSID).

Studying Panel 3a, notice that the mean earning change of formal sector stayers is approximately zero. This contrasts with the -0.04 we documented in Figure 2.2b, that focused on all agents who were initially in the formal sector. Also, the skewness is now -0.71 log-points, lower than the original -1.08. These two features indicate that the negative mean and high negative skewness are attributable to transitions out of formal employment. In an analogous fashion, when we compare Panel 3b and the results from the previous section, we can conclude that a large degree of the upside risk faced by informal sector workers is attributable to the possibility of finding a formal job. When we restrict attention to informal sector stayers (in contrast to all agents who were originally informally employed) the mean earning change is -0.04 (as opposed to zero) and displays some negative skewness (-0.11, as opposed to -0.02).

While one can already perceive that current employment status shape the income risk faced in the Brazilian labor market in important ways, a large degree of heterogeneity remains unexplored by the current analysis. For instance, are younger agents more likely to suffer sector transitions that will ultimately determine their long-term prospects? Are better paid workers more likely to suffer large negative drops in earnings? How transient is informality and how largely is it a phenomenon experienced by low-pay workers? To answer questions like these and further separate risk from heterogeneity, we study, in the next two subsections, transitions and the behavior of earning changes conditioning on current sector, earnings level, and age.

3.2 Transition across Sectors

As seen in Figure 2, large discrepancies between formal and informal sector workers' earnings are observed in the data. Worker characteristics might play an important role. Also, while the previous section focused on individuals that stay in the same form of labor arrangement across years, it is important to understand how much income risk is associated with transitions in employment status. In this section, we tackle both of these issues: first, how does sector transition risk behave as a function of age and earnings quantile? And second, what is the typical behavior of earning changes for agents that transition across the formal and informal sectors?

Figure 4 plots the yearly intensity of labor market sector transitions as a function of age (which

Figure 3: Earnings Innovations, stayers

Notes: The panel on the left displays the density of earnings innovations for formal sector stayers (agents that are observed in the same sector in two observations separated by one year). The panel on the right replicates this illustration for informal sector stayers. For a comparison of shapes, both panels also display a normal distribution with the same mean and variance.

varies across panels A through D) and deciles of the conditional earnings distributions (which vary along the horizontal axis). First, across all ages, better paid formal workers are slightly more likely to remain in the formal sector than worse paid workers, but the variation is quantitatively small. Second, better paid informal workers are always relatively more likely to transition to formal employment than lower paid workers. Third, transitions out of informality are quite frequent for younger workers (especially higher income, at almost 40% over a year), but tend to fade out for older workers.

Figure 5 studies the shape of log-earnings innovations for workers that are subject to a sectoral transition. In both of these figures, we aggregate over age, for simplicity (in the quantitative analysis of Section 4, we condition on age and current earnings). Workers transitioning toward the informal sector have, on average, a very large earning loss of -0.28 log-points. This distribution is also left skewed, with a skewness statistic of -0.74. A fairly opposite picture emerges for agents making the inverse transition: a mean earning gain of 0.18 log-points and a right-skewness of 0.51. Despite this almost reflected behavior in terms of average and skewness of earnings innovations, agents experiencing transitions face similar standard deviations of innovations (0.7 for agents transitioning into informality and 0.67 for agents transitioning away from it) and relatively high kurtosis (6.17 and 7.60, respectively).

Figure 4: Transition Frequencies

Notes: The four panels display average sector transition frequencies for agents in different age groups. I denotes informal employment, while F denotes formal employment. F-I denotes an agent transitioning from formal at t to informal employment at $t + 1$, while I-F denotes the opposite direction.

Figure 5: Earnings Innovations, transitions

Notes: The panel on the left displays the density of the earnings innovations for agents that are observed in formal employment at t , while in informal employment at $t + 1$. The panel on the right displays the density of innovations for agents that experienced the transition in the opposite direction. For a comparison of shapes, both panels also display a normal distribution with the same mean and variance.

3.3 More on key moments

For a deeper understanding of earnings risk and the consequences from heterogeneity, we disaggregate the four central moments not only by employment sector, but also by age and income groups following Guvenen et al. (2016). A first lesson from this exercise is that age has little quantitative impact on the patterns displayed for conditional means, standard deviation, skewness, or kurtosis. This does not imply that age is not an important determinant of labor earnings risk, as we have documented that it has consequences through the probability of transitions across the two forms of employment. We take advantage of the modest impact of age on conditional moments and present here simple plots, that disaggregate only by sectors and their conditional earning quantiles. Appendix B presents figures with data conditional on age.

Figure 6 plots mean log-earnings changes against earnings deciles (conditional on the original sector of employment). Notice that all curves in both panels are downward slopping, indicating mean reversion in earnings. On the left-hand side panel, we notice that this mean reversion is stronger for agents originally in informal employment arrangements, as noticeable by a steeper slope of the curve for the informal sector, and stronger at extreme quantiles, with a meaningful departure from a linear behavior. The right-hand side decomposes mean earning changes for sectoral stayers and individuals that were subject to sector transitions. A few additional features emerge. First, agents moving from formal

Figure 6: Mean

employment into an informal arrangement one year later are subject to negative mean earning shocks. This aligns with the view that this form of transition is essentially a negative shock, probably driven by an unwilling separation and a difficulty in obtaining an equivalent job. On the other hand, workers making transitions from informal employment to a formal arrangement are typically subject to large earning gains, especially among lower-income individuals.¹¹ The behavior of the average earning change for sector stayers is not markedly different from what we see in Panel 6a.

Figure 7 plots the standard deviation of log-earnings growth by conditional earnings decile. On Panel 7a, we see two important patterns emerge. First, there is more volatility around both extremes of the earnings distribution than around the median level. This U-shaped pattern is present in both sectors and survives further disaggregation. By inspecting Panel 7b, we also notice that workers that the terminal sector is the major determinant of realized labor earnings volatility (as measured by squared residuals). Individuals that move from the formal sector to the informal are typically the ones subject to the largest (typically negative) income innovations, while informal sector stayers are also subject to fairly risk earnings innovations. Formal sector stayers have the least volatile earnings. Overall, values between 0.4 and 0.9 for one-year variations of log-income are close in magnitude to those found in Guvenen et al. (2016) for the U.S. labor market, using tax-return data.

In Figure 8, we display the conditional skewness of earning changes. On Panel 8a, we see a typical left skewness, which is present in other contexts, and is stronger for higher deciles of the earnings

¹¹Notice that we see an income loss at the highest quantiles. This might be a compensation for additional benefits of formal sector employment, such as pension contributions and improved job stability.

Figure 7: Standard Deviation

Figure 8: Skewness

distribution. The possibility of frequent large negative shocks is a driver of precautionary savings of high earnings workers and fosters wealth inequality, as we study in Section 4. On Panel 8b, some nuances emerge. We can notice a relevant positive skewness for workers transitioning from informality to formal employment, as well as in lower deciles of the distribution of formal sector stayers. The former emphasize the importance of formalization shocks, that are not only typically positive on average, but are also asymmetric. Positive skewness does not find any counterpart in the American data, as studied in Guvenen et al. (2016).

Last, Figure 9 plots the conditional kurtosis of log-earning changes. As age, conditional earnings deciles do not play a large role in the determination of kurtosis, while the sectors of the labor market

Figure 9: Kurtosis

matter greatly. Formal sector workers display a high conditional kurtosis (15.71 when aggregated over ages and income deciles, while 21.06 when further conditioning on stayers), which is particularly extreme for low-earnings workers. These values are comparable to what was found for U.S. workers (Güvenen et al., 2016). Although higher than the normal distribution, earning changes for other sectoral transitions display more moderate kurtosis values. We conjecture that the asymmetry might be due to the larger degree of inflation indexing in formal labor contracts.

4 Quantitative Model

We now describe the model used to assess how the different earnings dynamics affect consumption and savings. We build the simplest setup able to capture the first-order effects. The tool is a partial equilibrium version of a life cycle model with heterogeneous agents in the tradition of Aiyagari (1994), İmrohoroglu, İmrohoroglu, and Joines (1995), and Huggett (1996). We follow the setup and calibration in De Nardi et al. (2016) quite closely.

Time is discrete and a model period is one year. Each agent is described by four variables: age j , sector s , current earnings y , and current assets a . They live with certainty until age J and then die. Individuals work until age $R < J$ and live as retirees until the end of their lives. Earnings are exogenous and evolve according to a sector and age-dependent first-order Markov process.

Conditional on being at age j , working in sector s , and with current earnings y , the probability of working in sector s' and having earnings y' in the next period is given by $\tilde{p}_j(s, y, s', y')$. This probability

can be decomposed into the product of two terms: the probability that the agent will switch sectors, $\gamma_j(s, y, s')$, and the conditional probability of moving to income group y' in the (potentially) new sector s' , $p_j(s, y, s', y')$.

Asset holdings are composed by a one-period riskless asset that pays an exogenous rate of return r . The agent may short the asset up to a limit ϕ . Households choose consumption c and next period's assets a' . They enjoy utility over consumption and maximize the discounted expected utility throughout life. The period utility function is $u(c) = \log c$ and the intertemporal discount factor is β .

The state vector of a typical working-age agent is (j, s, y, a) and its value function given by $v^w(j, s, y, a)$. The agents' problem can be recursively defined as:

$$v^w(j, s, y, a) = \max_{c, a'} \left\{ u(c) + \beta \sum_{y'} \sum_{s'} v(j+1, s', y', a') \gamma_j(s, y, s') p_j(s, y, s', y') \right\}$$

subject to

$$c + a' = y + (1 + r)a,$$

$$c > 0, a' \geq -\phi.$$

The continuation value function $v(j, s, y, a)$ depends on the agent's age and whether she is retired or not and is given by:

$$v(j, s, y, a) = \begin{cases} v^w(j, s, y, a) & \text{if } j < R \\ v^r(j, a) & \text{if } j \geq R. \end{cases}$$

A retired agent does not work, has zero income, and essentially solves a cake-eating problem, given the assets she accumulated over her working life. The problem can then be written recursively as follows:

$$v^r(j, a) = \max_{c, a'} \left\{ u(c) + \beta v^r(j+1, a') \right\}$$

subject to

$$c + a' = (1 + r)a,$$

$$c > 0, a' \geq -\phi, a_{J+1} \geq 0.$$

Note that in the last life period ($j = J$) agents consume all their resources, since they will be dead in the next period. Thus, the problem can be solved by backward induction from the last age to the first.

4.1 Calibration

The key to our quantitative exercise is the estimation of the exogenous earnings process. The usual approach to estimating it follows two steps. First, it is assumed a parametric AR(1) specification for earnings with normally distributed innovations. Then, a discretization method is applied to convert the continuous autoregressive process into a discrete first-order Markov chain. Although widely used, this method assumes that earnings growth follows a normal distribution, an assumption rejected by our data. Therefore, we deviate from it and apply a non-parametric method similar to the one used in Hansen, Hsu, and Lee (2014), De Nardi et al. (2016), and Ferreira and Gomes (2017). The novelty of this approach is that it makes no assumption about the distribution of earnings growth.¹²

The first step is to select a finite number of earnings observations to compose the earnings grid used in the simulation of the model. We decided for a grid that is sector and age dependent. Due to data limitation, we divided our sample into eight age groups: 18–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, and 55–65. Then, for each sector and age group, we choose ten points corresponding to the average earnings of each decile of the earnings distribution. The survey sample weights were used to compute the averages. These values represent the support of the transition probabilities described next.

The earnings transition probabilities are age dependent in our model. Then, for each age group, constructing the transition probabilities is done by simply calculating the fractions of workers who made the transition between sector-decile pairs in one year. This is only possible because of the panel structure of our data. Again, the survey sample weights were used to calculate the fractions. Note that there are twenty possible combinations of sector-decile pairs for each age group. Therefore, we end up estimating 3,200 transition probabilities. In the model, we repeat the same set of probabilities to all agents belonging to the same age group.

The remaining two parameters of the model were set according to recurrent values in the literature. The interest rate r was set to 0.02 and the intertemporal discount factor β to 0.96. This creates a wedge between the rate of time preference and the return on savings which is characteristic of incomplete market models.

¹²An alternative procedure was developed by Guvenen et al. (2016). They propose a parametric stochastic process for earnings that is more flexible than the usual approach and is able to capture the non-normality of the data. However, we chose the non-parametric method due to its simplicity and ease of implementation.

Figure 10: Consumption Dynamics

4.2 Results

4.2.1 Baseline

We start by documenting the key facts about implied consumption and wealth. Figure 10 plots moments of the consumption dynamics. On the left panel (10a), we see the evolution of mean consumption. Consistent with their higher income, agents in the formal sector consume more on average than agents in the informal one. Along the life-cycle, the path of consumption is increasing and we see a fanning out of the inequality between means of these groups. In other words, the average agent in the formal sector sees faster consumption growth than in the informal sector. The right panel (10b) plots the evolution of the variance of consumption. Some features invite attention. First, inequality is initially larger within the informal than within the formal sector. However, as within group dispersion rises faster among formal workers, both sectors stabilize at roughly the same level toward the end of the career. Overall consumption dispersion, which also incorporates inequality across groups, seems to behave as an upward translation of formal sector inequality.

Figure 11 plots the evolution of wealth. On the left panel, we see the evolution of per-capita wealth for each of the two groups. Agents in the formal sector accumulate more assets throughout their lives. The difference in asset holdings between groups roughly stabilizes when agents are in their 30s. To understand asset accumulation, it is important to analyze precautionary motives. First, for agents that remain in the same sector, earnings innovations are much safer among formal sector workers. As compared to informal workers, this points toward lower precautionary savings. However, transitions across sectors are also frequent. For agents in the formal sector, these transitions are left-skewed negative shocks, while for agents in the informal sector they have the opposite behavior. This force should

Figure 11: Wealth Dynamics

increase the precautionary savings of formal workers as compared to informal ones. Which of the two forces dominates is ultimately an empirical question, which is analyzed in Panel 11b. Except very early in their worklives (when transitions are very frequent), the former force dominates. Informal sector workers accumulate more assets relative to their incomes to smooth out (within sector) earnings risk.

4.2.2 The importance of data on informality

We use a large sample Brazilian survey (PNADC), for a country in which highly detailed social security data (RAIS) is available and has been frequently used in research. That is driven by a special concern for both: (i) capturing a large fraction of the labor market (approximately 40%) that is different and would be missed in social security data, (ii) understanding the importance of transitions in and out of formal employment. To describe what wrong conclusions would be obtained by using only RAIS data, we propose the following counter-factual exercise.

In the counter-factual, all agents are treated as being employed in the formal sector in all years of their lives. In the model, the probability of a transition out of formality is artificially set to zero. Earning shocks between years t and $t + 1$ are therefore given only by the conditional distribution for formal sector stayers p .

The first panel in Figure 12 plots consumption dynamics. First, the counterfactual exercises, illustrates that social security data would lead to underestimation of precautionary savings. Implied consumption is larger for formal workers in the counterfactual. Naturally, as mean income would be mismeasured mean consumption in the cross section would also be overestimated. Notice, by comparing the blue and gray lines, that the slope of the age-consumption profile would also be overestimated.

The second panel illustrates the evolution of the wealth-income ratio. The two lines for formal

Figure 12: Counter-factual exercise: ignoring informality

workers in the baseline exercise and the counterfactual are mostly parallel, up to the late 50s, when precautionary reasons vanish as the end of the career approaches. This is one side of the underestimation of precautionary motives, the one that originates from ignoring the downside risk of losing formal employment. A second source of underestimation is clear when comparing the mean wealth-income ratio in the baseline and the counterfactual. As we have discussed in the previous section, informal workers have an even larger precautionary motive for saving, so ignoring their existence leads to an additional bias coming from the composition of the sample studied.

4.2.3 The importance of non-parametric methods

TBW. What goes wrong when we use a simple AR(1)?

5 Conclusion

FIXME:TBW

A Residualized net earnings: Construction

Here we describe the procedure to calculate the earnings measure used throughout the article. Our starting point is the monthly gross income from the main job. The next step is to subtract contributions to Social Security. We applied the official rules of the *Instituto Nacional do Seguro Social* (INSS), which is the Brazilian Social Security Institute. We took into account that discount schedules vary by year and are not homogeneous across all workers. In Brazil, for instance, formal workers of the private and public sectors are subject to different Social Security regimes, and therefore face different marginal rates

of contribution. Also, some informal workers in our sample report contributing to Social Security on their own. In such cases, we applied the official rules for autonomous contributors. Next, we deduct income tax from the monthly earnings net of Social Security payments. We used the tax rules of the *Secretaria da Receita Federal do Brasil* (RFB), which is the Brazilian Internal Revenue Service. The tax schedules also vary by year but are homogeneous across all formal workers. On the other hand, informal workers are not subject to income taxes.

At this point, we have a monthly net income measure in nominal terms. Real earnings were calculated in 2018Q3 Reais (R\$) using the Brazilian price index *Índice Nacional de Preços ao Consumidor Amplo* (IPCA). We used the monthly and regional indexes provided by IBGE, taking into account regional differences in the inflation rate. To get an annualized measure, we multiply the monthly real net income of formal and informal workers by 13.33 and 12, respectively. The unusual multiplier adopted for formal workers has its roots in the regulation of the Brazilian formal sector. By law, such workers have entitled to an additional thirteenth salary every year plus one-third of wages due to paid vacation.

Finally, we remove age and temporal effects from the annualized real net earnings measure. We do so by regressing the logarithm of such measure on dummies of age and quarter. Formally, let $y_{i,j,t}$ be the annualized real net earnings of worker i , with age j , in quarter t . Then, the specification estimated by OLS is given by

$$\ln y_{i,j,t} = \delta_j + \gamma_t + e_{i,j,t},$$

where δ_j is the age dummy, γ_t is the quarter dummy, and $e_{i,j,t}$ is the residual. We make no assumptions about the structure of the residual, consistent with our premise of doing a non-parametric analysis. The estimated residual $\hat{e}_{i,j,t}$ is what we call the residual net earnings measure. Then, the one year change in the residual net earnings is defined as

$$\Delta \hat{e}_{i,j,t} \equiv \hat{e}_{i,j,t+4} - \hat{e}_{i,j,t}.$$

B Age and conditional moments: the relative unimportance of age

[Fixme: Complete]

Figure 13: Mean

Figure 14: Standard Deviation

Figure 15: Skewness

Figure 16: Kurtosis

References

- AIYAGARI, S. R. (1994): “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 109, 659–684.
- ALVAREZ, J. (2004): “Dynamics and Seasonality in Quarterly Panel Data: An Analysis of Earnings Mobility in Spain,” *Journal of Business & Economic Statistics*, 22, 443–456.
- ARELLANO, M., R. BLUNDELL, AND S. BONHOMME (2017): “Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework,” *Econometrica*, 85, 693–734.
- BAKER, M. AND G. SOLON (2003): “Earnings Dynamics and Inequality among Canadian Men, 1976–1992: Evidence from Longitudinal Income Tax Records,” *Journal of Labor Economics*, 21, 289–321.
- BOSCH, M. AND J. ESTEBAN-PRETEL (2012): “Job creation and job destruction in the presence of informal markets,” *Journal of Development Economics*, 98, 270 – 286.
- BOTELHO, F. AND V. PONCZEK (2011): “Segmentation in the Brazilian Labor Market,” *Economic Development and Cultural Change*, 59, 437–463.
- CAPPELLARI, L. (2004): “The Dynamics and Inequality of Italian Men’s Earnings: Long-Term Changes or Transitory Fluctuations?” *Journal of Human Resources*, 39, 475–499.
- CONSTANTINIDES, G. M. AND D. DUFFIE (1996): “Asset Pricing with Heterogeneous Consumers,” *Journal of Political Economy*, 104, 219–240.
- COŞAR, A. K., N. GUNER, AND J. TYBOUT (2016): “Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy,” *American Economic Review*, 106, 625–63.
- DE NARDI, M., G. FELLA, AND G. P. PARDO (2016): “The Implications of Richer Earnings Dynamics for Consumption and Wealth,” *Working Paper*.
- DE PAULA, A. AND J. A. SCHEINKMAN (2011): “The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil,” *Review of Income and Wealth*, 57, S8–S26.
- DICKENS, R. (2000): “The Evolution of Individual Male Earnings in Great Britain: 1975–95,” *The Economic Journal*, 110, 27–49.
- FERREIRA, P. C. AND D. B. P. GOMES (2017): “Health care reform or more affordable health care?” *Journal of Economic Dynamics and Control*, 79, 126–153.
- GUVENEN, F., F. KARAHAN, S. OZKAN, AND J. SONG (2016): “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?” *Working Paper*.
- GUVENEN, F., S. OZKAN, AND J. SONG (2014): “The Nature of Countercyclical Income Risk,” *Journal*

- of *Political Economy*, 122, 621–660.
- HANSEN, G. D., M. HSU, AND J. LEE (2014): “Health insurance reform: The impact of a medicare buy-in,” *Journal of Economic Dynamics and Control*, 45, 315–329.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2010): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118, 681–722.
- HUGGETT, M. (1993): “The risk-free rate in heterogeneous-agent incomplete-insurance economies,” *Journal of Economic Dynamics and Control*, 17, 953–969.
- (1996): “Wealth distribution in life-cycle economies,” *Journal of Monetary Economics*, 38, 469–494.
- İMROHOROGLU, A., S. İMROHOROGLU, AND D. H. JOINES (1995): “A life cycle analysis of social security,” *Economic Theory*, 6, 83–114.
- INTERNATIONAL LABOR ORGANIZATION (2013): “Women and Men in the Informal Economy: A Statistical Picture,” Second Edition.
- KREBS, T. (2003): “Human Capital Risk and Economic Growth,” *The Quarterly Journal of Economics*, 118, 709–744.
- KREBS, T. AND Y. YAO (2016): “Labor Market Risk in Germany,” *Working Paper*.
- KRUEGER, D. AND F. PERRI (2006): “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *The Review of Economic Studies*, 73, 163–193.
- MEGHIR, C., R. NARITA, AND J.-M. ROBIN (2015): “Wages and Informality in Developing Countries,” *American Economic Review*, 105, 1509–46.
- MEGHIR, C. AND L. PISTAFERRI (2011): “Earnings, Consumption and Life Cycle Choices,” *Handbook of Labor Economics*, 4, 773–854.
- ORDÓÑEZ, J. C. L. (2014): “Tax collection, the informal sector, and productivity,” *Review of Economic Dynamics*, 17, 262 – 286.
- PORTA, R. L. AND A. SHLEIFER (2008): “The Unofficial Economy and Economic Development,” *Brookings Papers on Economic Activity*, 39, 275–363.
- ROCHA, R., G. ULYSSEA, AND L. RACHTER (2018): “Do lower taxes reduce informality? Evidence from Brazil,” *Journal of Development Economics*, 134, 28 – 49.
- ULYSSEA, G. (2017): “Firms, Informality and Development: Theory and Evidence from Brazil,” *American Economic Review* (forthcoming).