Self-employment and Labor Market Risks

Richard Audoly*

This version: March 2021
First version: September 2018

Abstract

I study the labor market risks associated with being self-employed. I show that the self-employed are subject to larger labor income fluctuations than employees and that they frequently transition into unemployment. Given the self-employed are not eligible to unemployment insurance, I analyze the provision of benefits targeted at these risks using a calibrated search model with (i) precautionary savings, (ii) work opportunities in paid- and self-employment, (iii) skill heterogeneity. This exercise suggests that extending the current US unemployment insurance scheme to the self-employed does not improve the welfare of the low earners, while making their unemployment spells longer.

*Department of Economics, Norwegian School of Economics (NHH), `richard.audoly@gmail.com`.
This paper is a revised version of the second chapter of my doctoral dissertation. I am grateful to my supervisors at University College London, Fabien Postel-Vinay and Jan Eeckhout, for their continual guidance. For extremely valuable feedback, I wish to thank Rory McGee and Ludo Visschers, as well as my examiners, Rasmus Lentz and Wei Cui. I am also thankful to workshop and seminar participants at Aarhus (Sandbjerg Manor), the EEA 2020 Online Conference, UCL, and NHH for their comments.

1 Introduction

Job loss ranks amongst the most significant risks that workers face over the course of their career. Many countries target sizable transfers to the unemployed in the form of unemployment insurance (UI), and a substantial literature aims at characterizing the optimal UI contract.[1] By contrast, while the self-employed account on average for fifteen percent of employment across OECD countries, there is little evidence on the labor market risks associated with self-employment.[2] And, because the majority of the self-employed are not eligible to UI in these countries, traditional social insurance programs are not well-designed to alleviate these risks.[3]

This paper studies the labor market risks associated with self-employment and the provision of benefits targeted at these risks. My first contribution is to quantify the labor market risks faced by the self-employed in US survey data. My focus is on working-age individuals who are the main earners in their household. In each month, these workers are categorized as paid-employed if they derive most of their earnings from a job and as self-employed if they derive most of their earnings from a business. I show that (i) earnings are substantially more volatile in self-employment than in paid-employment; (ii) the self-employed frequently become unemployed, using a standard definition of unemployment.

My second contribution is to build a framework to analyze several benefit schemes targeted at these risks. I develop and calibrate a search model with precautionary savings and use it to quantify the welfare changes and distortions in labor supply implied by the introduction of these policies.

A key pre-requisite to my calibration is to allow for substantial worker heterogeneity in the model. This heterogeneity matters because the data show that self-employed workers are over-represented in the tails of the earnings distribution, which suggests the

[1]Following Jacobson et al. (1993), many studies have confirmed that job loss has long- term negative effects on workers' earnings. Examples of work on the optimal UI contract include Acemoglu and Shimer (1999), Chetty (2008), and Kolsrud et al. (2018).

[2]Throughout the paper, I use the terms "self-employed" and "self-employment" to designate all workers who get most of their labor income from a business. This definition is further clarified when I introduce the data in Section 2.

[3]Among these countries, a handful offer some form of public unemployment insurance for some narrow group of self-employed workers, such as artists and writers in Germany. These schemes are reviewed in details in OECD (2018). In the US, some owners of incorporated businesses (S- and C-corps) can become eligible under very specific conditions. Sole proprietors, partnership-owned business, Limited Liability Corporations, and independent contractors do not qualify. This restriction was abolished as part of the US government response to Covid-19 (CARES Act).

underlying skills differ widely across workers. To discipline this feature of the model, I follow Bonhomme et al. (2019) and use a k-means algorithm to partition workers based on their observed labor income and likelihood to become unemployed. These two measures are strongly correlated in the resulting clusters of workers. For instance, relative to the high earnings group in self-employment, the low earners are close to five time as likely to make a transition from self-employment to unemployment. Building heterogeneity in such a way into the model allows to study how self-employed workers at different skill levels respond to the introduction of additional benefit entitlements.

The model is calibrated to replicate the empirical evidence on the exposure of the self-employed to labor market risks. Paid-employment and self-employment each come with specific labor income and unemployment shocks. In line with the data, the model captures the substantial flows between paid-employment, self-employment, and unemployment, acknowledging that moving between paid- and self-employment can also represent a response to unemployment shocks or bad realizations of earnings. In addition, workers can also partially insure against labor market risks by borrowing and drawing down their savings, as well as by relying on household-level income (spousal income and welfare transfers). My calibration matches the substantial fraction of households at lower wealth levels, a well-known feature of the data that also applies to the self-employed.

I use the calibrated model to study several policies designed to alleviate the labor market risks associated with self-employment. I center on two broad classes of policies: a UI scheme for the self-employed and a basic income scheme. This exercise suggests that effectively improving the welfare of the self-employed at the bottom of the earnings distribution requires some degree of progressivity in financing these benefits. My results indicate that extending the UI system currently in place in the US, including the cap on income used as a basis to compute unemployment tax liabilities, has barely any impact on the welfare of low earners, while it leads them to spend more time in unemployment. Self-employed workers close to the middle of the earnings distribution actually value this type of policies more than the group at the bottom. By contrast, a modest basic income scheme does a good job at targeting the low earnings group and reduces the time they spend unemployed. At the top of the earnings distribution, however, this type of unconditional transfer results in substantial welfare losses for high earners, who bear the burden of financing the policy.

More generally, the analysis developed in this paper is relevant for the design of policy in several dimensions. It first shows that, much like regular employees, a fraction of the self-employed are not insulated from labor market risks. Second, the rise of alternative work arrangements with the emergence of labor platforms (firms that match workers to customers without being bound to them by an employment contract) may further increase the number of self-employed with low earnings in the future. My framework highlights some of the key trade-offs in providing additional benefits to this group of workers.

Related literature. This work is related to the large literature studying the risks faced by wage workers in the labor market, most notably unemployment, and the associated optimal provision of unemployment insurance (Chetty, 2008; Kolsrud et al., 2018). My approach is closest to the series of papers that study UI benefits within the context of a fully specified structural model (Acemoglu and Shimer, 1999; Lentz, 2009; Krusell et al., 2010). I depart from these studies by allowing agents to find work both in paid-employment and self-employment and by focusing on the labor market risks specifically associated with being self-employed.

This work also contributes to the growing literature on self-employment. Most of these papers are concerned with the factors leading workers to becoming self-employed (Hamilton, 2000; Hurst and Lusardi, 2004; Levine and Rubinstein, 2017; Humphries, 2017; Catherine, 2019; Jones and Pratap, 2020). This line of research does not directly study the risks associated with self-employment in the labor market. Perhaps the closest paper to mine in that regard is Catherine (2019), who quantifies the value of paid-employment as a back up option for workers deciding whether to start a business, though he does not consider the risk of becoming unemployed.

A branch of the self-employment literature also studies to what extent self-employment can be a path out of unemployment by analyzing the outcomes of the ensuing businesses (Hombert et al., 2020; Camarero Garcia and Murmann, 2020). While this channel is present in my framework, I instead center on labor market risks conditional on workers being self-employed. My framework allows me to assess the impact of several policies targeted at these risks.

Outline. The next section presents the data. Section 3 introduces the model. Section 4 discusses the calibration procedure. Section 5 gives results on several counter-factual policies designed to provide additional insurance to the self-employed, and section 6 concludes.

2 Self-employment and labor market risks in the data

This section provides empirical evidence on the labor market risks associated with self-employment. Studying these risks in the data requires frequent records of each individual's labor market history to capture transitions in and out of unemployment, information on wealth, which represents a key mean to self-insure, and a large enough sample, since the self-employed represent a small fraction of total employment. The Survey of Income and Program Participation meets these requirements (Census Bureau, 2014, SIPP thereafter). I pool together the four panels spanning 1996-2013. I stress two key dimensions of labor market risks in these data: labor income changes while employed (an intensive margin risk) and unemployment (an extensive margin risk). I start by showing how individuals with at least some experience of self-employment over the sample period differ from those always in paid-employment. I then document the exposure of the self-employed to labor market risks.

2.1 The self-employed in the SIPP

The sample is restricted to the working-age individuals with the largest earnings in each household over the duration of the survey. This restriction is made both to center on the individuals with the strongest ties to the labor market and to acknowledge that at least some insurance against labor market risks takes place at the household level, through shared assets, labor income from other household members, and household-level welfare programs. I incorporate these additional sources of insurance in the model introduced in Section 3.

Workers are assigned to one of three labor market states in each month: paid-employed (P), self-employed (S), or unemployed (U). When in employment, they are categorized as paid- or self-employed based on their primary source of labor income. I choose to focus on earnings as hours appear to be imprecisely measured in the SIPP. Earnings also more directly relate to the resources workers can actually set aside in anticipation of labor market shocks. For an overwhelming majority of workers, this definition clearly singles out one labor form (paid- or self-employment) as an individual's main source of labor income (see figure C.2). I use a standard definition of unemployment, based on whether workers actively search for work over the duration of a non-employment spell. Appendix A provides a complete description of the data

	Employees	Self-employed
Age	39.864	42.575
Gender (woman=1)	0.415	0.263
Married (married=1)	0.539	0.616
Race (non-white=1)	0.196	0.138
HS Graduate	0.902	0.894
College Graduate	0.454	0.474
Post-graduate	0.099	0.130

Table 1: Demographic characteristics by subsample. See main text and Appendix A for data sources and definitions.

and the definition of each labor market state.

Defining the self-employed as workers deriving most of their income from their business encompasses individuals working in very different ways, both in terms of scale and industry. It groups together freelancers, contractors, doctors and lawyers, as well as owner-managers with potentially many employees. By definition, however, they all have in common that they are not the employee of a distinct person (strictly speaking some of them can receive a wage from their business). The model developed in subsequent sections allows for permanent worker heterogeneity in skills to capture the variety of situations covered by this definition.

Given this definition of self-employment, I further cut the sample of main earners into two subsamples: The self-employed sample, which is made of workers who are self-employed at some point during the survey, and the employee sample, which is made of workers always working as employees over the duration of the survey. The self-employed sample represents roughly fifteen percent of the sample of main earners.[4]

Table 1 gives some basic demographic statistics on these two groups of workers. The self-employed are more likely to be older men, married, and not to belong to a minority. These two groups are otherwise broadly similar in terms of education level, with the self-employed slightly more likely to hold a post-graduate qualification.

Table 2 further compares the self-employed and employee samples in terms of income and wealth. All monetary values are given in 2009 real dollars throughout. The self-employed are characterized by more heterogeneity along these dimensions than

[4]This is higher than the share of self-employment in total employment implied by the data from the Bureau of Labor Statistics. The sample is, however, selected on labor market attachment. Figure C.1a shows the implied self-employment rate derived from the SIPP is very close to the BLS series when taking all working age individuals into account.

	Sample	Percentile p10	p25	p50	p75	p90
Labor income (monthly, $2009)						
Main earner	Employee	1,341	2,093	3,296	5,033	7,328
	Self-employed	857	1,635	3,065	5,539	9,406
Other earners	Employee	0	0	0	2,180	3,861
	Self-employed	0	0	0	1,892	3,780
Liquid wealth ($2009)						
Net liquid wealth	Employee	-12,572	-885	3,296	49,859	185,564
	Self-employed	-16,254	-1,323	2,992	67,000	256,079
Unsecured debt	Employee	0	0	1,472	8,887	23,983
	Self-employed	0	0	1,408	9,834	27,693

Table 2: Earnings and wealth: self-employed vs employee sample. See main text and Appendix A for data sources and definitions.

employees. Both the lowest and highest earners are self-employed. For instance, the tenth percentile of monthly labor income is $857 ($1,311) and the ninetieth percentile is $9,406 ($7,328), respectively for workers with (without) self-employment. The distribution of labor income from other household members look, to the contrary, similar. Turning to wealth measures, a substantial fraction of the self-employed have not accumulated more wealth than employees to self-insure. The median net liquid wealth (net worth excluding business, home, and vehicle equity, aggregated at the household level) is approximately the same for the self-employed and employees (around $3,000). The self-employed also do not seem to have more access to credit, as the distribution of unsecured debt is similar across samples. Taken together, these statistics suggest that, in terms of household finance, a substantial fraction of the self-employed is potentially no less insulated from labor market risks than employees.

The data further shows that most of the businesses from which self-employment income is derived are relatively limited operations.[5] In the self-employed sample, the reported median business assets for workers in self-employment are slightly above $6,000. Only 32 percent of the corresponding businesses are incorporated, and only 5 percent of the self-employed declare that their business has ever had more than 25 employees since it was established. These figures suggest that the scope to use one's business as a

[5]This is also consistent with the finding that most business owners are not "entrepreneurs." Hurst and Pugsley (2011), for example, find that most new business owners do not intend to grow or bring new products to the market.

From/To	P	S	U
P	–	0.022	0.024
S	0.008	–	0.005
U	0.141	0.094	–

Table 3: Monthly transition rates across labor force states. See main text and Appendix A for data sources and definitions.

saving vehicle, and therefore as an additional source of wealth that can be drawn upon if needed, is limited for a substantial fraction of the self-employed.[6]

2.2 Exposure to labor market risk in the self-employed sample

I restrict the analysis to the self-employed sample and document their exposure to labor market risks. These empirical regularities are the moment targets used to calibrate the model introduced in section 3.

Table 3 first shows that there are large worker flows between paid-employment, self-employment, and unemployment. In a typical month, there is a 2.2 percent chance that a paid-employed worker makes a direct transition to self-employment (without an unemployment spell in between). Workers also appear to exit unemployment at a faster rate as employees (from U to P) than as self-employed (from U to S). In terms of transitions to unemployment, these workers do make direct transitions from self-employment to unemployment (from S to U). While the probability to become unemployed after a spell as a wage worker (from P to U) is larger, workers in the self-employed sample are on average markedly more likely to experience a PU transition: the PU rate in the employee sample is .011. Using this group as a benchmark, the workers in the self-employed sample then make a SU transition at half the rate of regular wage workers. There is therefore a non-trivial flow of workers going self-employment to unemployment.

Table 4 reports summary statistics on labor income and wealth for the self-employed sample. Within this sample, earnings also show more dispersion in self-employment than in paid-employment. The 10th percentile of labor income is almost $200 lower in self-employment. It appears that workers start their own activity or maintain them

[6] There is evidence that this is a possibility for some self-employed. Miller et al. (2019) find that UK owner-managers shift income from their business over time to avoid being taxed at higher marginal rates. See table C.1 for additional data on business wealth.

Worker current employment status	Percentile p10	p25	p50	p75	p90
Labor income ($2009, monthly)					
Paid-employed (P)	974	1,674	2,907	4,792	7,722
Self-employed (S)	790	1,605	3,144	5,889	10,081
Labor income growth within employment type (6 months continuously employed)					
Paid-employed (P)	-0.315	-0.045	-0.008	0.114	0.541
Self-employed (S)	-0.561	-0.219	-0.008	0.320	1.341
Labor income growth between employment type (at least 6 months, direct transition)					
PS	-0.707	-0.400	-0.028	0.424	1.255
SP	-0.579	-0.221	0.218	0.974	2.635
Net liquid wealth ($2009)					
Paid-employed (P)	-17,208	-2,518	753	40,276	193,287
Self-employed (S)	-15,459	-764	5,818	83,850	287,921

Table 4: Earnings and wealth by labor force status in the self-employed sample. See main text and Appendix A for data sources and definitions.

with lower associated labor income than as an employee. In addition, conditional on being continuously employed in the same labor form, earnings are substantially more volatile in self-employment. On top of the risk of becoming unemployed, self-employment also comes with large fluctuations in labor income. Over a six-month period, for instance, there is a 25% chance that one's earnings drop by more than 20% in self-employment. The corresponding drop in earnings for workers in paid-employment is 4%.

Table 4 also shows that workers tend to leave self-employment for a clearly better paying job (the median growth following a SP is about twenty percent), while PS transitions are associated with a stagnation in earnings (median growth close to zero). As many of these transitions are likely to be voluntary, as by definition there is no intervening unemployment spell, this pattern is consistent with the existence of non-pecuniary benefits in self-employment (Hamilton, 2000; Catherine, 2019). In terms of liquid wealth holdings finally, a substantial fraction of these workers do not have access to large savings to self-insure against unemployment and fluctuations in earnings. Even for workers in self-employment, who mostly are not eligible to UI benefits, more than 25 percent have negative liquid wealth.

Summary. Three main points emerge from these descriptive statistics with regard to the self-employed's exposure to labor market risks. First, if some self-employed rank amongst the highest earners, many of them are also at the lower end of the earnings distribution. Second, there are some clear labor market risks associated with self-employment. The self-employed do become unemployed at a non-trivial rate, and they experience higher labor income volatility than in paid-employment. Third, many of them have limited wealth reserves that they can draw on in the face of these events. In the next sections, I present a quantitative model that can account for these empirical regularities.

3 A search model with self-employment and savings

I build a search model in which risk-averse workers can save and borrow to jointly capture (i) workers' transitions between paid-employment, self-employment, and unemployment; (ii) fluctuations in labor income; (iii) a self-insurance motive. This framework relates to several studies that describe a frictional labor market where workers can move across several forms of employment, such as formal and informal employment (Meghir et al., 2015) or public sector and private sector employment (Bradley et al., 2017). I depart from these prior studies by making agents risk-averse in a partial insurance framework.[7] Agents can self-insure against labor market risks by borrowing or drawing down their savings, as well as by considering additional sources of income at the household level, such as spousal earnings and welfare transfers.

3.1 Environment

Time is discrete. The labor force is represented by a continuum of working age individuals with measure one. These workers are the model counterpart to the "main earners" defined in section 2. They are risk-averse and discount the future at rate $\beta < 1$. Their per-period utility of consumption $c > 0$ is given by utility function u.[8] Workers are allowed to borrow and save using a risk-free asset, a, with rate of return r and borrowing limit $\underline{a} \leq 0$ exogenously given.

[7] See also Lise (2013) for a search model with savings in which workers climb a single job ladder.
[8] $u : \mathbb{R}_+^* \to \mathbb{R}$ is assumed to satisfy $u' > 0$, $u'' < 0$, and $\lim_{c \to 0} u'(c) = \infty$.

Workers can be in one of four labor market states: paid-employment (P), self-employment (S), unemployed on UI benefits (B), or unemployed not eligible to benefits (U). They can search for work opportunities either as paid- or self-employed when unemployed. While employed, they can only search for opportunities in the alternative employment form.[9] In the baseline model, workers in unemployment are only eligible to UI benefits if they were previously employees. Self-employed workers terminating their business are not eligible to such transfers.

Workers differ in skills, which translate in permanent differences in earnings potential. I index this heterogeneity by $k = 1, \ldots, K$. A worker's skill level conditions the distributions of business and job opportunities from which they draw. These distributions are denoted by F_k^P and F_k^S, respectively for workers of type k drawing a wage or "self-employment income." There is no recall of past jobs or business opportunities. In addition, workers experience fluctuations in labor income while working at a job or at their business. The income processes governing these shocks, conditional on current labor income, also depend on k and are given by $Q_k^P(.|w)$ and $Q_k^S(.|y)$.

A worker's skill level also conditions the destruction rate of employment opportunities. Jobs and businesses disappear, respectively, with exogenous probability δ_k^P and δ_k^S. These destruction rates are indexed on ability k to allow for differential exposure to unemployment risk by worker type. In addition, workers are always free to leave their current job or business, in which case they become unemployed without access to unemployment benefits (they go straight to labor market state $s = U$).

Unemployment insurance (UI) has two dimensions in the model: b^P and T^P. b^P is a benefit function, mapping workers' wage in their last job to some benefit level. T^P controls the potential duration of benefits payments, with the eligibility to benefits expiring at rate $1/T^P$. These assumptions restrict UI to belong to the "constant benefit, finite duration" class of UI policies, in line with the US system.

There are additional income sources that accrue to the household beyond the main earner's labor income and the (potentially negative) returns to wealth ra. There may be additional earners in the household, and, on top of unemployment insurance, other welfare programs represent additional income. Household income is modeled as a tuple of functions, $\left\{Y_k^s,\ s \in \{U, P, S\}\right\}$, of the main earner's labor income in the corresponding labor market state s and worker type k. These functions give a reduced-form representation of the phasing-out of welfare programs as earnings increase and of

[9] SS and PP transitions are captured by earnings shocks.

added worker effects.

3.2 Timing

Each period t unfolds as follows:

1. Earning realization. Workers in paid- and self-employment get a new labor income draw, respectively from $Q_k^P(.|w)$ and $Q_k^S(.|y)$.

2. Quits and separations. Conditional on earnings in the present period, workers can decide to quit, in which case they become unemployed with no access to UI benefits (U). If not, they are hit by an exogenous destruction shock with probability $\delta_k^s,\ s \in \{P, S\}$.

3. Search. Workers not separated, as well as those previously in unemployment sample job offers and self-employment opportunities, respectively, with probabilities $\lambda^{sP},\ s \in \{U, S\}$ and $\lambda^{sS},\ s \in \{U, P\}$. These probabilities are assumed to be mutually exclusive, so that unemployed workers, who can search for work opportunities both as paid- or self-employed, get at most one labor income draw in each period, either from F_k^P or F_k^S. If they choose to pursue this job or business opportunity, they immediately switch to this new labor form and earn the associated wage or business income.

4. Consumption and savings. Household income accrues to all workers. UI benefits are paid out to the eligible fraction of unemployed workers. Agents then choose consumption c and next period's net wealth a'.

3.3 Worker's problem

Notations. Let $R_k^s(a, y)$ be the present value of being in state s with net wealth holdings a and labor income y for a worker of type k at the start of the quits and separations stage. Let $V_k^s(a, y)$ stand for the worker's present value at the start of the consumption and savings stage. I denote the net gain of getting a job or business opportunity (a draw from $F_k^{s'}$) as

$$\phi_k^{ss'}(a, y) := \int \max\Big\{V_k^{s'}(a, \tilde{y}) - V_k^s(a, y), 0\Big\} dF_k^{s'}(\tilde{y})$$

where s is the worker's current state, and $s' \in \{P, S\}$, which by assumption is only defined for $s \neq s'$.

Value functions: paid-employed. The value of holding a job with current wage w and wealth a at the beginning of the period is given by

$$R_k^P(a,w) = \max\left\{V_k^U(a),\ \delta_k^P V_k^B(a,w) + (1-\delta_k^P)\Big[V_k^P(a,w) + \lambda^{PS}\phi_k^{PS}(a,w)\Big]\right\}. \quad (1)$$

Recall that only layoffs (δ_k^P-shocks) entitle workers to unemployment benefits. Voluntary quits (the first term after the max operator) result in unemployment without benefits. The value of being in paid-employment with wage w and wealth a at the consumption and savings stage writes

$$V_k^P(a,w) = \max_{c,\tilde{a}}\left\{u(c) + \beta\int R_k^P(\tilde{a},\tilde{w})dQ_k^P(\tilde{w}|w)\right\} \quad (2)$$
$$\text{s.t.}\quad c + \frac{\tilde{a}}{1+r} = Y_k^P(w) + a;\quad \tilde{a} \geq \underline{a}.$$

Value functions: self-employed. The value of a business at the beginning of the period mirrors that of job holders in equation (1), with the difference that unemployment transitions are all to the no-benefits state (U):

$$R_k^S(a,y) = \max\left\{V_k^U(a),\ \delta_k^S V_k^U(a) + (1-\delta_k^S)\Big[V_k^S(a,y) + \lambda^{SP}\phi_k^{SP}(a,y)\Big]\right\}. \quad (3)$$

The self-employed's savings problem is given by

$$V_k^S(a,y) = \max_{c,\tilde{a}}\left\{u(\kappa\cdot c) + \beta\int R_k^S(\tilde{a},\tilde{y})dQ_k^S(\tilde{y}|y)\right\} \quad (4)$$
$$\text{s.t.}\quad c + \frac{\tilde{a}}{1+r} = Y_k^S(y) + a;\quad \tilde{a} \geq \underline{a}.$$

The key difference with the savings problem of the paid-employed is that self-employment comes with non-pecuniary benefits, a factor $\kappa \geq 1$, scaling consumption directly in the utility function. These non-pecuniary benefits capture a preference for being self-employed, which could stem from the intrinsic value of being one's own boss, from unreported earnings, or from other tax advantages.[10] Empirically, this feature is motivated in the data by the non-trivial growth in earnings observed following a direct SP transition, which is interpreted as the compensating differential of self-employment. Non-pecuniary benefits can potentially lead workers to turn down a job offering a better pay than what they get from their current business.[11]

Value functions: unemployed. The value of searching for unemployed workers on benefits is given by

$$R_k^B(a,w) = \lambda^{UP}\phi_k^{BP}(a,w) + \lambda^{US}\phi_k^{BS}(a,w), \tag{5}$$

as workers in unemployment can search both for work as an employee and business opportunities. The corresponding savings problem is given by

$$\begin{aligned} V_k^B(a,w) &= \max_{c,\tilde{a}} \left\{ u(c) + \beta \left[\left(1 - \frac{1}{T}\right) \cdot R_k^B(\tilde{a},w) + \frac{1}{T} \cdot R_k^U(\tilde{a}) \right] \right\} \\ \text{s.t.} \quad c + \frac{\tilde{a}}{1+r} &= Y_k^U + a + b(w); \quad \tilde{a} \geq \underline{a}, \end{aligned} \tag{6}$$

where the term in squared brackets after the discount factor comes from benefits expiring at rate $1/T$. The value functions for unemployed workers not eligible to benefits (state $s = U$) follow directly from adapting equations (5) and (6). They are given by

$$R_k^U(a) = \lambda^{UP}\mu_k^{UP}(a) + \lambda^{US}\mu_k^{US}(a), \tag{7}$$

and

$$\begin{aligned} V_k^U(a) &= \max_{c,\tilde{a}} \left\{ u(c) + \beta R_k^U(\tilde{a}) \right\} \\ \text{s.t.} \quad c + \frac{\tilde{a}}{1+r} &= Y_k^U + a; \quad \tilde{a} \geq \underline{a}. \end{aligned} \tag{8}$$

3.4 Stationary equilibrium

Taken together, the optimal choices of consumption and savings, as well as the reservation strategies implied by workers' decisions to quit and to follow through on alterna-

[10]Private consumption through one's business is an example of such a tax advantage.

[11]Most structural models of self-employment allow for some degree of preference for business ownership (Humphries, 2017; Catherine, 2019; Jones and Pratap, 2020).

tive employment opportunities imply a stationary distribution over employment states $\{U, B, S, P\}$, labor income (y, w), and net wealth (a) for each skill level k. These distributions are denoted Γ_k in what follows. A formal statement of this definition can be found in Appendix B.

3.5 Discussion

I conclude the exposition of the model by briefly discussing some of my modeling choices. First, I do not model investment and the associated accumulation of business capital. Hurst and Lusardi (2004) argue that liquidity constraints are unlikely to be binding for a majority of entrepreneurs in the US, which would imply that rental markets for capital function well. This is also consistent with the fact that many businesses operate with a low capital stock. Median business assets in the SIPP are less than $6,500, suggesting that many self-employed have limited capital requirements. For the subset of workers in self-employment operating on a larger scale, an interesting extension would consider how these workers allocate their portfolio between liquid assets, their business, and other illiquid assets, given the labor market risks associated with this employment form.

In addition, I abstract from general equilibrium effects. The self-employed account for less than 15 percent of workers in the sample. Their borrowing and saving behavior can therefore be expected to have a very limited impact on the aggregate interest rate. Besides, while my definition of self-employment covers the majority of firm owners, most of the businesses associated with self-employment have very few, if any, employees. The data suggest that only 5 percent of these businesses have ever had more than 25 employees. Large firms account for the majority of US employment,[12] and labor market risks are unlikely to be a first-order concern for the self-employed at these firms. It seems therefore reasonable to assume that aggregate labor market outcomes are largely orthogonal to the introduction of additional benefits targeted at the self-employed.

Lastly, I model worker skills k as unidimensional. As such, I implicitly assume that these skills are transferable across employment forms. By focusing on data from the self-employed sample in the calibration, I ensure that workers are skilled enough in self-employment that it becomes their main source of labor income at some point. Because many self-employed in the data are not observed working as employees, recovering a

[12]Data from Census Bureau show that 57 percent of US employment is at firms with more than 100 workers.

joint distribution of skills in both labor forms is not feasible. The life cycle income profiles reported in Humphries (2017) for various groups of self-employed suggest that experience as a wage worker is, to a large extent, transferable to self-employment, conditional on making the transition.

4 Calibration

My calibration strategy proceeds in two steps. Worker heterogeneity is first introduced by clustering workers with similar labor market outcomes. In a second step, I target key moments from the data to calibrate the remaining parameters, conditional on the heterogeneity uncovered in the first step. Data from the self-employed sample (workers with at least some self-employment over the duration of the survey) are used throughout.

4.1 Worker heterogeneity

To discipline worker heterogeneity in the model (k), I rely on a clustering tool from machine learning to partition workers into distinct groups.[13] This classification aims at capturing the sizable heterogeneity in earnings found in the self-employed sample (see section 2).

Clustering algorithm. Following Bonhomme et al. (2019), I use a k-means algorithm to group workers in earnings classes in the SIPP data. This procedure finds the best partition of the data according to the following objective function

$$\underset{\tilde{h},k_1,\ldots,k_N}{\arg\min} \sum_{i=1}^{N} \left\| \hat{h}_i - \tilde{h}(k_i) \right\|^2, \tag{9}$$

where N is sample size, $k_i \in \{1, ..., K\}$ are partitions of $\{1, ..., N\}$ with $1 < K \leq N$, $\hat{h}_i$ is a vector of features used to group workers, and $\tilde{h}(k_i)$ is the corresponding vector of features for group k to which i is assigned. Each element in $\tilde{h}(k_i)$ is computed by averaging over the members of the group. The solution to the minimization in (9) then assigns a cluster to each i such that the squared Euclidean distance between i's vector

[13] Thereafter, the terms “class”, “group”, and “cluster” are used interchangeably to designate the outcome of the clustering procedure.

of characteristics and the average of these characteristics in i's group is minimized.[14]

The logic behind this clustering step is in essence the same as that described in Bonhomme et al. (2019). They first cluster firms based on their empirical distribution of earnings amongst employed workers, before using the resulting classes in a series of mixture models. The estimated partition then captures both observed and unobserved firm heterogeneity.

In my framework, I rely on a similar partition of the data to discipline worker heterogeneity in earnings. In the model, this partition then translates into cluster-specific distributions of employment opportunities $\left\{F_k^P, F_k^S\right\}$ and unemployment shocks $\left\{\delta_k^P, \delta_k^S\right\}$.

Implementation. To implement this procedure in practice, one needs to choose the number of clusters, K, and the vector of features $\hat{h}_i$. I set $K = 4$, which offers a good trade-off between capturing a reasonable degree of heterogeneity in worker ability and computation time.[15]

The vector of features $\hat{h}_i$ includes two key outcomes: a measure of labor income when employed and a measure of exposure to unemployment. The former is the empirical cdf of earnings

$$\widehat{ECDF}_i(y) := \frac{1}{\sum_t \mathbb{1}\{\text{Employment}_{it} = 1\}} \sum_t \mathbb{1}\{y_{it} \leq y\},$$

with y_{it} denoting labor income for i in period t and Employment_{it} the corresponding indicator for employment. This distribution is based on all earnings, irrespective of whether they come from paid- or self-employment at a point in time.[16] It is computed for each decile of the empirical distribution of labor income in the sample.

Exposure to unemployment is the fraction of time workers spend in non-employment over the duration of the survey. Letting T_i be the number of months the individual is observed in the panel (typically slightly under three years in the 1996-2008 SIPP

[14] Standard algorithms to efficiently solve this global minimization problem are readily available in standard packages. I use the implementation of the "Hartigan-Wong" algorithm in the base R function "kmeans".

[15] Figure C.3 shows how the total within sum of square changes with K.

[16] I do not further distinguish between labor income from paid- versus self-employment since many workers are never observed in paid-employment in my sample.

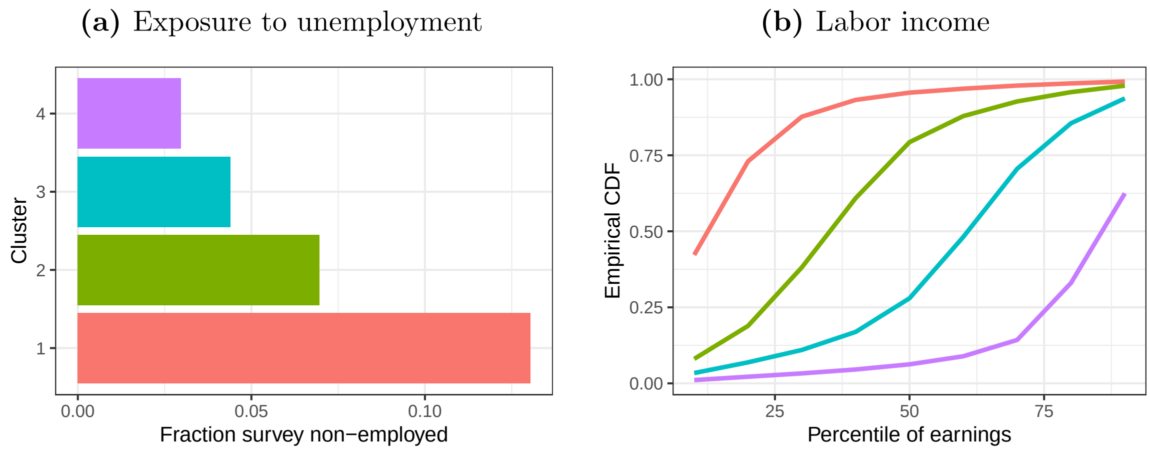


Figure 1: Worker classes obtained from clustering algorithm. The figure depicts the average feature in each cluster with $K = 4$. The labels are ordered based on the value of the empirical cdf of labor income at the median.

panels), it is given by

$$\widehat{LMA}_i := \frac{1}{T_i} \sum_t \mathbb{1}\{\text{Employment}_{it} = 0\}.$$

Results. The resulting average features for each worker class, $\tilde{h}(k_i)$, are depicted in figure 1. The figure shows a clear relationship between the two measures used to build the clusters: the group with the largest labor income is also least likely to be unemployed. This negative relationship between earnings when employed and time spent in unemployment holds for all worker groups.

Table 5 further shows how worker classes differ along several dimensions not directly used in the clustering procedure. The four groups are ordered based on the median labor income in each group. In terms of demographic composition, workers with lower earnings are more likely to be women and belong to a minority, and are less likely to be married and have a college education. (Recall that the sample is restricted to the main earner in each household.) Reassuringly, age is roughly the same in all groups on average, so the classification is not driven by workers' accumulation of human capital. The resulting clusters of workers also clearly relate to the large variation in the scale of the corresponding businesses in terms of equity, employment, and legal status. For example, the group of highest earners are close to three times as likely

	Worker class k (median earnings)			
	\$1,049	\$2,230	\$3,830	\$7,494
Number of workers (N_k)	3,515	5,366	5,373	3,825
Demographic characteristics				
Age	42.44	41.45	42.63	44.20
Gender (woman=1)	0.47	0.28	0.19	0.15
Married (married=1)	0.42	0.57	0.68	0.76
Race (non-white=1)	0.19	0.15	0.12	0.10
College Graduate	0.45	0.47	0.47	0.51
Post-graduate	0.07	0.09	0.13	0.24
Business characteristics				
Median business equity (\$2009)	476	2,458	8,819	22,658
Incorporated	0.17	0.27	0.36	0.46
More than 25 employees	0.02	0.03	0.05	0.10

Table 5: Characteristics of worker clusters. See main text and Appendix A for data sources and definitions.

to be incorporated (46%) as the group of lowest earners (17%).[17] The worker classes resulting from the algorithm therefore capture meaningful observed characteristics, as well as potentially unobserved ability traits determining earnings.

4.2 Model parametrization

UI benefits. The actual unemployment insurance system in the US is determined at the state level. I follow Chetty (2008) and Saporta-Eksten (2014) and approximate these policies as a fifty percent replacement rate for a maximum duration of six months. I also cap the maximum monthly benefit payments at \$2,000.[18] In the notation of the model, the UI system is then given by the replacement rate function $b^P(w) = \max\left\{2,000; .5 \cdot w\right\}$ and $T^P = 6$. The relevant model wage w is assumed to be the last wage realization before the worker becomes unemployed.

[17] I also report the distribution of worker groups by industry in figure C.4.

[18] This further approximates the caps on UI benefit payments enacted by most states. See the "Significant Provisions Of State Unemployment Insurance Laws" tables from the Department of Labor for the corresponding weekly ceilings by state.

Household income functions. I obtain a mapping between an individual's current labor income and household income by estimating the following regressions

$$\ln Y_{it}^{s} = \alpha_{k}^{s} + \beta_{k}^{s} \cdot \ln y_{it}^{s} + \epsilon_{it},$$

where Y_{it}^{s} is total household income, including other earners and welfare transfers, but net of UI payments, $\ln y_{it}^{s}$ is the main earner's labor income, and ϵ_{it} is an error term. The index $k = 1, \ldots, K$ refers to the worker's assigned cluster, and $s \in \{P, S, U\}$ denotes their current state in the labor market. (The slope β_{k}^{s} is omitted for $s = U$, as labor income is zero in this case by definition.) The estimated coefficients $\left(\widehat{\alpha}_{k}^{s}, \widehat{\beta}_{k}^{s}\right)$ give the corresponding household income function by cluster and labor force status. These coefficients are reported in table C.2.

Utility and labor income distributions. Utility from consumption is given by a standard CRRA specification $u(c) = (1-\gamma)^{-1} c^{1-\gamma}$. The distribution of labor income draws, F_{k}^{s}, is assumed to follow a log-normal distribution with mean μ_{k}^{s} and standard deviation σ_{k}^{s} for each $k \in \{1, \ldots, K\}$ and $s \in \{P, S\}$. Shocks to labor income are modeled as a chance $1 - \rho^{s}$ to redraw from the corresponding F_{k}^{s} distribution in each period.[19]

4.3 Calibration procedure

I set the returns on net liquid wealth r to zero and the coefficient of risk-aversion γ to 3/2. The lower bound on net liquid wealth is allowed to vary by worker group according to $\underline{a}_{k} = \underline{a} \cdot Y_{k}^{\mathrm{p50}}$, where Y_{k}^{p50} denotes median household income in a cluster and the factor $\underline{a}$ is taken from Kaplan and Violante (2014).[20] The borrowing constraint is therefore tighter for workers with lower earnings potential.

The remaining parameters are calibrated by matching a set of moments from the data to their counterpart in the simulated model. Starting from the stationary distribution of workers in each group, Γ_{k}, I simulate a panel of N_{k} workers (the number of observations in each cluster k in the data) for 48 months. I stack all cluster-specific panels and obtain the model simulated moments by using the exact same definitions

[19] This income process better fits the distribution of labor income growth within labor form, which exhibits a clear peak at zero, than a standard AR(1) process.

[20] These authors set this factor to the median ratio of households' "total credit limit" to their quarterly income, which is equal to 74% in the Survey of Consumer Finances. I then set $\underline{a} = -.74 \times 3$.

as those used to construct the moments derived from the data.

Though all parameters are jointly calibrated, I motivate my choice of moment targets by mapping each set of calibrated parameters to specific moments. The job and business arrival rates $\left(\lambda^{US}, \lambda^{UP}, \lambda^{PS}, \lambda^{SP}\right)$ are chosen to replicate the corresponding monthly transition rates. The job and business destruction rates $\left(\delta_k^P, \delta_k^S\right)$ directly relate to the k-specific transition rates into unemployment, respectively from paid- and self-employment. The parameters governing the distributions of labor income draws $\left(\mu_k^s, \sigma_k^s\right)$ are set to match the observed distributions of income by worker group k and employment state s. The rates of persistence of income $\left(\rho^P, \rho^S\right)$ are disciplined by the six-month growth of income for workers continuously employed in paid- and self-employment, respectively. The discount factor $\left(\beta\right)$ is set to match the median net liquid wealth in the data. Finally, the non-pecuniary benefits of self-employment $\left(\kappa\right)$ is controlled by the gap in median labor income growth following a direct PS and SP transition, where labor income growth is computed only for transitions with at least six months of continuous employment history on either side of the transition.

The classification of workers is implemented by relating class-specific parameters to class-specific targets. In line with my choice of clustering variables (unemployment risk and labor income) I therefore make the job and business destruction parameters and labor income parameters class-specific.[21] I stress that, while there is a tight link between these parameters and the corresponding moments, they still need to be calibrated jointly due to the selection across P, S, and U, that naturally arises in the model. As an example, the model simulated distribution of labor income in self-employment is shaped by workers' reservation income strategy with respect to paid-employment and unemployment given their current wealth level.

4.4 Results

I report the model fit to the targeted moments in tables 6 and 7. Table 6 first shows the fit to the moments that are specific to each worker class.[22] Both for PU and SU transitions, there is a clear decreasing pattern where higher earners, defined by median earnings in the class, are much less likely to transition to unemployment. There is an almost one percent chance for the lowest earners to make a SU transition in a

[21] The UP and US transition rates, which also relate to unemployment, do not exhibit a clear pattern by class in the data, so the corresponding arrival rates are common to all workers.

[22] The corresponding class-specific parameters are given in table C.3.

given month, while the highest earners face a .2 percent chance of a similar event. The class-specific destruction parameters allow the model to very closely replicate this heterogeneity in the chance to make a transition to unemployment.

The labor income distributions also show that, even within a worker class, labor income tends to be markedly more dispersed in self-employment than in paid-employment. While median earnings are very similar for both types of employment in most worker classes, self-employment income tends to have a longer left tails for the lowest earners. Overall, the model replicates these distribution well, though for some worker classes it does not fully get the tails right given the log-normal assumption.

Table 7 shows the model fit to the moment computed on the whole sample, and table 8 lists the parameters common to all workers. The model replicates well the large flows across unemployment, paid-, and self-employment observed in the data. In terms of the arrival rates of job and business opportunities for employed workers (λ^{SP} and λ^{PS}), the model sees transitions to paid-employment as relatively straightforward for self-employed workers (a six percent chance in every month), though many of these opportunities are turned down (the actual transition rate is slightly under one percent). The model also emulates the difference in earnings volatility observed within employment type found in the data, with workers in self-employment experiencing markedly larger changes in earnings over a six-month period of continuous employment.

The median growth rate of earnings following an employment-to-employment transition differs starkly between SP and PS transitions. The calibrated non-pecuniary benefits of self-employment, κ, implies that consumption is 2.5% more valuable in self-employment for these workers. The model can replicate a large fraction of the gap in median labor income growth between the two types of transitions found in the data.

Regarding net liquid wealth, finally, the model implies median net liquid wealth holdings in line with the data, though, in spite of having access to UI benefits in the event they become unemployed, workers in paid-employment still tend to accumulate slightly more wealth than in the data. I also stress that despite having access to some insurance through additional welfare programs and the income of other household members in unemployment (captured by Y_k^U) and despite setting the coefficient of relative risk-aversion at a relatively low value ($\gamma = 3/2$), the yearly discount rate required to match the fraction of workers at low wealth levels is high (approximately 0.10). The gap between the discount rate and the interest rate is implicitly the cost of using savings as insurance, and the gap required to match median net liquid wealth

	Worker class k (median \$2009 earnings)							
	1,032		2,227		3,840		7,626	
Moment	Model	Actual	Model	Actual	Model	Actual	Model	Actual
Transition rates to U								
PU	0.041	0.044	0.029	0.028	0.018	0.018	0.016	0.015
SU	0.012	0.009	0.006	0.006	0.004	0.004	0.002	0.002
Earnings in P (log)								
p10	6.743	6.170	7.161	7.078	7.813	7.718	8.441	8.305
p25	6.882	6.609	7.452	7.410	8.010	7.988	8.676	8.644
p50	7.022	7.041	7.744	7.721	8.245	8.262	8.958	8.941
p75	7.189	7.366	7.987	8.035	8.480	8.541	9.287	9.231
p90	7.328	7.663	8.278	8.277	8.716	8.807	9.522	9.523
Earnings in S (log)								
p10	5.993	5.924	6.967	6.769	7.545	7.353	8.340	8.136
p25	6.429	6.352	7.312	7.253	7.883	7.845	8.638	8.534
p50	6.952	6.894	7.656	7.695	8.221	8.246	8.936	8.939
p75	7.388	7.316	8.069	8.048	8.627	8.614	9.294	9.333
p90	7.824	7.803	8.414	8.461	8.965	8.981	9.592	9.592

Table 6: Model fit—moments specific to worker class.

Moment	Model	Actual
Transition rates		
UP	0.164	0.141
US	0.106	0.094
SP	0.007	0.008
PS	0.022	0.022
Earnings growth in paid-employment (6 months)		
p10	-0.385	-0.315
p25	-0.136	-0.045
p50	0.000	-0.008
p75	0.087	0.114
p90	0.518	0.541
Earnings growth in self-employment (6 months)		
p10	-0.582	-0.561
p25	-0.287	-0.219
p50	0.000	-0.008
p75	0.317	0.320
p90	1.250	1.341
Median earnings growth after direct transition		
After PS	0.058	-0.028
After SP	0.243	0.218
Fraction below actual median wealth ($2009)		
P (p50 in data: $ 753)	0.334	0.500
S (p50 in data: $5818)	0.461	0.500

Table 7: Model fit—moments computed on whole sample.

Parameter	Description	Value
Set exogenously		
r	Rate of return on liquid wealth	0.0
γ	Coefficient of rel. risk aversion	1.5
Calibrated		
β	Discount factor (monthly)	0.992
κ	$u(\kappa \cdot c)$ in self-employment	1.025
λ^{UP}	Arrival rate P opportunity from U	0.168
λ^{US}	Arrival rate S opportunity from U	0.111
λ^{SP}	Arrival rate P opportunity from S	0.067
λ^{PS}	Arrival rate S opportunity from P	0.030
ρ^P	Persistence income in P	0.818
ρ^S	Persistence income in S	0.775

Table 8: Parameters—Common across worker classes.

holdings in the data is large. Lentz (2009) notes that workers tend to value unemployment benefits more when this gap is larger in his structural model estimated on Danish data. Through the lens of the model, additional social insurance is therefore potentially quite valuable for some groups of self-employed.

5 Policy analysis: social insurance for the self-employed

What are the labor market and welfare effects of providing benefits targeted at the labor market risks associated with self-employment? This section assesses the impact of several potential policies through the lens of the calibrated model. I emphasize how heterogeneous groups of self-employed workers respond to these policies.

5.1 Alternative policies

UI for the self-employed. A first potential policy is to extend UI benefits to the self-employed. I rely on a similar formulation as the one used in the baseline model. The UI policy for the self-employed is given by $b^S(y) = \max\{b^S_R \cdot y, \bar{b}^S\}$: benefits are paid at a replacement rate b^S_R of workers' last income capped at $\bar{b}^S$ for a maximum duration of T^S months. These benefits are financed by taxing the labor income of workers in self-employment according to the tax schedule τ^S. Introducing some notation, τ^S must satisfy

$$\sum_k \omega_k \left[\int \tau^S(y) d\Gamma_k^S(a,y) - \int b^S(y) d\Gamma_k^C(a,y) \right] = 0, \tag{10}$$

where state $s = C$ denotes previously self-employed workers currently on benefits (the counterpart to state $s = B$), Γ_k^s denotes the measure of workers in labor force state s and worker class k, and $\omega_k := N_k/N$ is the share of workers in class k.[23]

When writing down the self-employed's problem, I assume that, contrary to the paid-employed, the self-employed can choose to terminate their activity and receive benefits. This is because the distinction between layoffs and voluntary quits, which conditions eligibility in many UI systems, is irrelevant for the self-employed. Formally, the value of self-employment at the beginning of the search stage (equation (3) in the baseline model) now reads

$$R_k^S(a,y) = \max \left\{ V_k^C(a,y),\ \delta_k^S V_k^C(a,y) + (1-\delta_k^S)\left[V_k^S(a,y) + \lambda^{SP} \phi_k^{SP}(a,y) \right] \right\},$$

where $V_k^C(a,y)$ is the value of becoming unemployed when benefits are paid out. There are then two ways for the self-employed to cash these benefits: either following a δ_k^S-shock (involuntary) or by "choosing" to become unemployed (voluntary).

Basic income policy. Because the activity of the self-employed at the bottom of the earnings distribution is potentially difficult to monitor, I also consider a basic income "top up" on existing transfers, irrespective of the worker's labor force status. Such a transfer is by design targeted at the workers with the lowest earnings.

This policy is implemented through the following tax function $\mathcal{T}(Y) := -b + \tau Y$, where b is some unconditional cash transfer and τ is a constant tax rate of household income Y. The parameters (b, τ) are assumed to be such that the government budget is balanced, or more formally,

$$\sum_{s,k} \omega_k \int \mathcal{T}\big(Y_k^s(y)\big) d\Gamma_k^s(a,y) = 0. \tag{11}$$

[23] In practice, this type of policy might generate non-negligible overhead costs to administer the program. It probably requires the self-employed to make their activity more formal in legal terms so that such taxes can be raised (the incorporation rate is 32% in the self-employed sample).

In implementing this policy scenario, it is assumed that the UI program for the paid-employed is discontinued and the corresponding resources added to the overall budget.

Assessing alternative policies. I study both the labor market impact and the welfare impact of these policies by comparing steady-states. Distortions in the labor supply of workers across labor force status are summarized by the shift in the transition rates between paid-employment, self-employment, and unemployment induced by the alternative policies.

I use two metrics for welfare. I first follow Krusell et al. (2010) and study the equivalent variation in consumption, $\mathrm{EV}_k^s(a, y)$, which is defined as

$$E_t \sum_{\tilde{t}=0}^{\infty} \beta^{\tilde{t}} u\Big(\big(1 + \mathrm{EV}_k^{s_t}(a_t, y_t)\big) c_{t+\tilde{t}}\Big) = E_t \sum_{\tilde{t}=0}^{\infty} \beta^{\tilde{t}} u(\hat{c}_{t+\tilde{t}}),$$

where $\hat{c}_t$ denotes consumption choices under one of the policies and c_t is consumption in the baseline economy. $\mathrm{EV}_k^s(a, y)$ is the permanent (potentially negative) change in consumption that makes the agent indifferent between the baseline and the economy with the policy, starting from a given employment state (s_t), income (y_t), and asset position (a_t). I also calculate the compensating grant, $\mathrm{CG}_k^s(a, y)$, defined as $V_k^s(a + \mathrm{CG}_k^s(a, y), y) = \tilde{V}_k^s(a, y)$. $\mathrm{CG}_k^s(a, y)$ is the cash transfer making the agent indifferent between the baseline economy (value function V_k^s) and the economy under an alternative policy (value function $\tilde{V}_k^s$). I report the average of these measures by worker class and employment state.

5.2 Results

I study three different policy scenarios, two alternative UI policies and one basic income policy. The first UI scenario (UI-A) simply aligns the policy parameters to those of the paid-employed. The replacement rate is set to fifty percent, the potential benefit duration to six months, and the cap on monthly payments to \$2,000 ($b_R^S = .5$, $T^S = 6$, and $\bar{b}^S = 2000$ in the notation introduced above). The tax schedule to finance these benefits is set to $\tau^S(y) = \min\{\tau \cdot y, 7000/12\}$. This formulation is chosen to emulate the way UI contributions are collected in the actual US system, since, in practice, all States cap the annual labor income subject to UI contributions.[24] Because UI contributions

[24] \$7,000 corresponds to the States with the lowest threshold. Detailed tables on UI contribution rates by State can be found at `https://oui.doleta.gov/unemploy/avg_employ.asp`.

are infra-marginal for all but the lowest earners in such a system, I also consider a more progressive schedule (UI-P) where UI contributions are paid out on all self-employment income at a constant rate $\tau^S(y) = \tau \cdot y$, without a cap. The benefit entitlements are kept identical to the first UI scenario. I finally study a basic income policy (BI), an unconditional cash transfer $b = 100$ financed by levying taxes at a constant rate τ of household income. The value of b is chosen to make the total size of the program (total government income) broadly consistent with that implied by the two UI policies.

These policies are implemented by finding the tax rate τ satisfying the budget balance conditions in equations (10) and (11). In addition, I allow these tax rates to have an effect along the intensive margin of labor supply in the UI-P and BI scenario. (UI contributions are essentially infra-marginal in the UI-A scenario.) I simply input an elasticity of labor income to the marginal tax rate of .5.[25] The policies and corresponding tax rates are summarized in table C.4.

Table 9 shows how employment choices are distorted in response to the introduction of each policy. The table gives the percentage change in each transition rate with respect to its baseline value. I contrast the changes in labor supply for the classes of workers at both end of the earnings distribution: the low earners (median monthly earnings of \$1,032) and high earners (median monthly earnings of \$7,626). UI benefits for the self-employed has a clear moral hazard component for the low earners. Their SU rate increases by close to twenty percent; their US rate drops by about two percent as well. For the low earners, capping UI contributions to finance UI payments implies that they shoulder more of the tax burden. As such, while self-employment becomes more attractive for workers currently in employment (decrease in SP, increase in PS), these changes are small. This is in contrast to the more progressive policy (UI-P), which makes self-employment a lot more attractive to the low earners, since they get the same insurance level with a lower tax burden. By comparison, the high earners do not respond strongly along the worker flow margin. The intuition is that they do not face strong liquidity constraints, and therefore their choice of employment state is not impacted by the availability of UI benefits in self-employment.

The basic income policy affects labor supply decisions in a very different way. The response in both group is broadly similar, though with different magnitudes. Overall, the combination of removing UI benefits for the paid-employed and of giving an unconditional cash transfer makes self-employment more attractive, especially with regard

[25] This value is an upper bound on the steady-state intensive margin labor supply elasticity in Chetty et al. (2011), who report .33 as their central estimate.

	Worker class (median \$2009)					
	1,032			7,626		
	UI-A	UI-P	BI	UI-A	UI-P	BI
Transition rates						
(% change wrt baseline)						
UP	-0.05	-0.70	-0.18	0.88	-0.23	1.08
US	-2.57	0.45	6.15	0.12	-0.10	0.84
PS	0.41	5.30	23.16	0.24	-0.39	2.22
SP	-1.35	-4.81	-18.67	-2.00	0.72	-18.38
PU	-0.23	-0.13	-0.17	2.03	2.52	1.14
SU	23.77	17.16	-7.61	-0.08	0.15	0.17
Share of labor force						
(% change wrt baseline)						
U	5.31	2.72	-10.40	-0.12	1.89	-6.38
P	1.10	-1.67	-12.80	-0.77	1.15	-14.89
S	-2.74	1.57	19.64	0.09	-0.15	1.74

Table 9: Change in transition rates in each policy scenario.

to the choices of workers already in employment. This is reflected in the shift in both SP and PS transitions. A key difference of this policy with respect to the UI policies is the absence of moral hazard effects, since the payment associated with the BI policy is not conditional on a worker's current labor force status. By boosting their earnings in employment and providing them with some degree of insurance in case they become unemployed, this policy strongly favors employment in the group, both along the SU margin (decrease by 13%) and US margin (increase by 7%). The high earners similarly tend to shift their labor supply toward self-employment, though these changes tend to be smaller.

Table 10 reports the welfare effects of these policies, broken down by worker class and employment state. The two measures of welfare, the equivalent variation in consumption (EV_k^s) and the compensating cash grant (CG_k^s, which is scaled by median earnings in the worker class to account for permanent heterogeneity in earnings), convey the same overall message in terms of which policy is more valuable to each worker class.

The results suggest that introducing a UI policy close to the current UI scheme for the paid-employed has a small, but mostly positive welfare impact. Across policies and groups, workers do not value the compensating cash grant at more than one month of median earnings overall. This result is quite striking given that, with a relatively high discount rate, precautionary savings are costly for these workers. The "align" policy scenario (UI-A) is the only policy considered that translates in positive welfare gains for all worker classes and employment states. Given the relatively low earnings disregard, the largest welfare gains are obtained for the medium-low and medium-high earnings groups (median monthly earnings $2,227 and $3,840, respectively). For the low earners (median monthly earnings $1,032), the welfare gains from having access to an extra source of insurance are in large part offset by the tax burden implied by the cap on taxable earnings. This comes on top of the increase in unemployment triggered by the policy in this group of workers documented in table 9.

Introducing UI in a more progressive fashion by removing the earnings disregard changes the distribution of welfare gains. The UI-P policy implies that workers in the low and medium-low earnings groups witness the largest welfare gains. These welfare gains have a very similar magnitude in both groups. However, in this policy scenario the high earnings group experiences a loss in welfare, since they are barely exposed to unemployment risks in the first place, hold more wealth to self-insure, and shoulder more of the tax burden once the cap on UI contributions is lifted.

The basic income policy, by far the most progressive policy design considered, translates into much larger welfare changes for the various groups of workers. For instance, the compensating cash grant implied by the BI policy is more than four times the median monthly labor income for the low earners. By comparison, the compensating cash grant implied by the UI-P policy for the low earners is more than four times smaller. On the other hand, these sizable welfare gains for the low earners come with large losses for workers in the medium-high and high earnings groups, who bear the burden of financing the policy. Using the same compensating cash grant metrics, the loss in welfare represents about two months of median earnings in the high earnings group.

To sum up, the policy analysis points to the existence of stark trade-offs in designing policies geared toward the risks faced by the self-employed in the labor market. Unconditional cash transfers are a potentially efficient tool to alleviate these risks for the low earners, and they do not discourage labor supply. However, this policy comes at the cost of sizable welfare losses for high earners. UI policies tend to have much more balanced welfare effects, even positive for all workers in the "align" scenario, but

	Worker class (median \$2009)											
	1,032			2,227			3,840			7,626		
	UI-A	UI-P	BI	UI-A	UI-P	BI	UI-A	UI-P	BI	UI-A	UI-P	BI
Equivalent variation in c – EV_k^s (percent)												
All s	0.04	0.25	1.57	0.21	0.33	0.39	0.20	0.16	-0.41	0.09	-0.08	-1.04
P	0.03	0.22	1.75	0.17	0.27	0.45	0.16	0.13	-0.37	0.07	-0.07	-1.02
S	0.03	0.28	1.74	0.20	0.33	0.46	0.19	0.16	-0.37	0.09	-0.09	-1.02
B	0.03	0.24	0.99	0.18	0.29	-0.62	0.18	0.15	-1.58	0.08	-0.07	-1.91
U	0.03	0.24	1.68	0.18	0.29	0.38	0.16	0.13	-0.42	0.06	-0.08	-1.00
C	0.56	0.79	—	1.22	1.33	—	1.39	1.37	—	1.37	1.22	—
Compensating grant in a – CG_k^s (ratio to median earnings)												
All s	0.10	0.71	4.64	0.45	0.71	0.89	0.40	0.33	-0.83	0.16	-0.16	-1.79
P	0.09	0.63	4.38	0.37	0.59	0.35	0.32	0.27	-1.21	0.13	-0.12	-2.03
S	0.08	0.79	5.10	0.46	0.74	1.05	0.41	0.33	-0.76	0.16	-0.17	-1.77
B	0.08	0.57	2.52	0.32	0.51	-0.84	0.25	0.21	-1.74	0.10	-0.08	-1.91
U	0.07	0.62	4.56	0.33	0.54	0.70	0.32	0.26	-0.82	0.13	-0.17	-1.96
C	1.24	1.76	—	1.81	2.00	—	1.50	1.46	—	0.92	0.76	—

Table 10: Welfare changes by worker class in response to each alternative policy.

are poorly targeted at the low earners. In addition, UI policies have clear moral hazard effects as they lead the low earners to spend more time unemployed.

6 Conclusion

In this paper, I study the labor market risks associated with self-employment. My first set of results are empirical. I show that many self-employed experience large income fluctuations and that there exist non-trivial self-employment to unemployment flows. In addition, many self-employed do not earn more or have more liquid wealth than employees, thus potentially limiting their resources to self-insure. These findings are for a sample of individuals who account for the majority of earnings within their household.

I then calibrate a search model with precautionary savings to replicate these empirical regularities. I use the model to assess how different groups of self-employed value several policies providing these workers with extra insurance against labor market risks. My second set of results point to the difficulty of improving the welfare of the self-employed with low earnings without a reasonable degree of progressivity in the design of these policies. Specifically, my analysis suggests that extending the UI system in place in the US to the self-employed results in (i) an increase in unemployment, (ii) very small welfare gains for the self-employed with low earnings. This type of framework can therefore inform the current policy debate on the rise of "gig economy" work and the degree to which this category of workers can potentially be reached by conventional social insurance programs.

The model in this paper offers a unified framework to study the labor market risks faced by all non-employees. In doing so, it abstracts from explicitly modeling some dimensions of running a business that are potentially relevant to some subset of the self-employed. An important direction for future work would be to understand how business owners allocate their portfolio between liquid assets, illiquid assets, and their business. Given the inherently risky nature of owning a business, such an extension would provide additional insights into the self-insurance dimension of entering and exiting self-employment.

References

Acemoglu, D. and R. Shimer (1999, October). Efficient Unemployment Insurance. *Journal of Political Economy 107*(5), 893–928.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A Distributional Framework for Matched Employer Employee Data. *Econometrica 87*(3), 699–739.

Bradley, J., F. Postel-Vinay, and H. Turon (2017). Public sector wage policy and labor market equilibrium: A structural model. *Journal of the European Economic Association*, jvw026.

Camarero Garcia, S. and M. Murmann (2020). Unemployment benefit duration and startup success. *ZEW-Centre for European Economic Research Discussion Paper* (20-033).

Catherine, S. (2019). Keeping options open: What motivates entrepreneurs? *Available at SSRN 3274879*.

Census Bureau (2014). Survey of Income and Program Participation. `https://www.census.gov/programs-surveys/sipp.html`. 1996, 2001, 2004, and 2008 panels.

Chetty, R. (2008, April). Moral Hazard versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy 116*(2), 173–234.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review 101*(3), 471–75.

Hamilton, B. H. (2000, June). Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy 108*(3), 604.

Hombert, J., A. Schoar, D. Sraer, and D. Thesmar (2020). Can unemployment insurance spur entrepreneurial activity? evidence from france. *The Journal of Finance 75*(3), 1247–1285.

Humphries, J. E. (2017). The causes and consequences of self-employment over the life cycle. *Yale University*.

Hurst, E. and A. Lusardi (2004, April). Liquidity Constraints, Household Wealth, and Entrepreneurship. *Journal of Political Economy 112*(2), 319–347.

Hurst, E. and B. W. Pugsley (2011, May). What Do Small Businesses Do? Working Paper 17041, National Bureau of Economic Research.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. *The American economic review*, 685–709.

Jones, J. B. and S. Pratap (2020). An estimated structural model of entrepreneurial behavior. *American Economic Review 110*(9), 2859–98.

Kaplan, G. and G. L. Violante (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica 82*(4), 1199–1239.

Kolsrud, J., C. Landais, P. Nilsson, and J. Spinnewijn (2018). The optimal timing of unemployment benefits: Theory and evidence from sweden. *American Economic Review 108*(4-5), 985–1033.

Krusell, P., T. Mukoyama, and A. Şahin (2010, October). Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations. *Review of Economic Studies 77*(4), 1477–1507.

Lentz, R. (2009, January). Optimal unemployment insurance in an estimated job search model with savings. *Review of Economic Dynamics 12*(1), 37–57.

Levine, R. and Y. Rubinstein (2017, May). Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More? *The Quarterly Journal of Economics 132*(2), 963–1018.

Lise, J. (2013, January). On-the-Job Search and Precautionary Savings. *The Review of Economic Studies 80*(3), 1086–1113.

Meghir, C., R. Narita, and J.-M. Robin (2015, April). Wages and Informality in Developing Countries†. *American Economic Review 105*(4), 1509–1546.

Miller, H., T. Pope, and K. Smith (2019). Intertemporal income shifting and the taxation of owner-managed businesses. Technical report, IFS Working Paper WP19/25, https://www. ifs. org. uk/publications/14475.

OECD (2018). *The Future of Social Protection.*

Saporta-Eksten, I. (2014). Job loss, consumption and unemployment insurance. *Unpublished manuscript*, 1–50.

A Data

The main data source is the Survey of Income and Program Participation (SIPP), a survey with detailed information on households' use of welfare programs (Census Bureau, 2014). I use data from the 1996, 2001, 2004, and 2008 panels. This Appendix expands on the definition of paid-employment and self-employment and describes the construction of the "main earners" and "self-employed" samples.

Labor force status. The SIPP contains two key pieces of information to classify individuals as unemployed, paid-, or self-employed: a week-by-week account of their employment state (employed, on layoff, unemployed, non-participating) and information on up to two jobs and two businesses (job/business identifier, wages, profits, incorporation status, and so on).

Employed individuals are classified as paid- or self-employed on the basis of the jobs or businesses for which they report most of their earnings. I start by cleaning the job and business identifiers provided in the survey, so that they are consistent with the start and end date reported for each job or business. For a large majority of workers, only one type of employment is reported: wage work or self-employment. In the case where workers report both working as an employee and having a business, they are assigned the status in which they report the largest earnings during the overlapping spell. In the event workers report more than one job or business, I simply add the corresponding wages/profits to get a measure of earnings in paid- or self-employment. I proceed similarly to define the business characteristics reported in the main text. For categorical variables (incorporation status and number of employees) I use the maximum across the two businesses.

The SIPP reports three types of non-employment week-by-week: on layoff, no job looking for work, no job not looking for work. To address potential differences in reporting between previously paid- and self-employed workers, I define as unemployment spells all non-employment spells of at most fifty weeks if the individual declares to be looking for a job at some point during the spell or if the spell ends with a transition to employment. The first part of the definition is standard (Chetty, 2008, Appendix B). The second part accounts for the fact former business owners might associate unemployment with the receipt of UI benefits.

Finally, I follow the convention in the Current Population Survey and build a monthly panel of employment state for each individual based on the second week (the first full week) in each month. Keeping only working-age individuals, I check the validity of my definitions by plotting the implied unemployment rate (figure C.1b) and self-employment rate (figure C.1a) against data from the Bureau of Labor Statistics (BLS).[26] The SIPP data are aggregated using cross-sectional weights. The figures show that my definitions yield sensible unemployment and self-employment rates.

Sample: Main earners. I retain individuals aged at least 25 or at most 65 across the survey. In each household, I only keep the main earner, defined as the individuals with the largest labor earnings across the survey, irrespective of their gender. I also exclude individuals who are mostly out of the labor force over the survey period. The aim of these restrictions is to reduce the sample to the individuals most strongly attached to the labor market. Taken together, these restrictions yield a dataset of 122,662 individuals which are followed on average for 35.5 months.

Sample: Self-employed. The key analysis and estimation sample is made of individuals who are part of the main earner sample and spending at least one month in self-employment over the duration of the panel. This additional restriction leaves me with a sample of 17,884 workers. All the statistics targeted in the calibration procedure are drawn from the self-employed sample.

B Equilibrium definition

I introduce some additional notation for workers' reservation strategies to formally describe the equilibrium. Let d_k^s be an indicator denoting workers' decision to quit their job or shut down their business given the realization of income. For instance, for a self-employed worker deciding to terminate their activity,

$$d_k^S(a,y) := \mathbb{1}\left\{V_k^U(a) > \delta_k^S V_k^U(a) + (1-\delta_k^S)\Big[V_k^S(a,y) + \lambda^{SP}\phi_k^{SP}(a,y)\Big]\right\}.$$

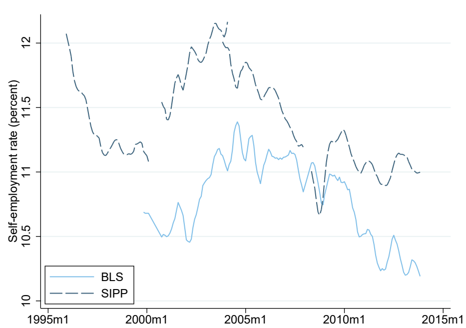
Let $\rho_k^{ss'}(a,y^s)$ be the reservation income draw from $F_k^{s'}$ making workers indifferent between their current labor force status s and employment in s'. This reservation income is implicitly given by $V_k^s\big(a,y^s\big) = V_k^{s'}\big(a,\rho_k^{ss'}(a,y^s)\big)$.

[26] The corresponding BLS series are LNU02048984 ("Incorporated Self-Employment"), LNU02027714 ("Non-incorporated Self-Employment"), LNU03000000 ("Unemployment"), and LNU02000000 ("Employment").

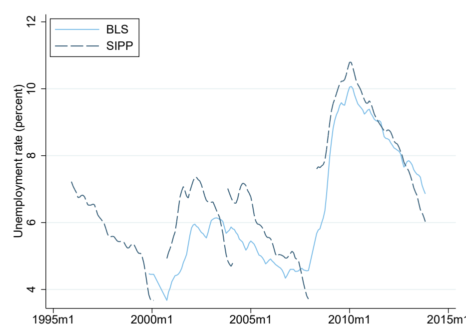
For each $k \in \{1, ..., K\}$ and $s \in \{U, B, P, S\}$, a stationary equilibrium is a set of value functions V_k^s and R_k^s, decision rules d_k^S, $\rho_k^{ss'}$, $\hat{a}_k^s$ and $\hat{c}_k^s$, and a distribution Γ_k across labor force states, wealth and income, such that

1. The R_k^s functions are defined by equations (1), (3), (5), and (7);
2. The choice to terminate employment, $d_k^S(a, y)$, and the reservation income functions $\rho_k^{ss'}(a, y^s)$ solve (1), (3), (5), and (7);
3. The V_k^s functions are defined by equations (2), (4), (6), (8);
4. The asset and consumption choice functions, $\hat{a}_k^s$ and $\hat{c}_k^s$, solve equations (2), (4), (6), (8);
5. Finally define $\boldsymbol{Q}_k$ the operator mapping the current distribution of workers to that in the next period. This operator arises from workers' optimal consumption and savings decisions, as well as their reservation strategies. The associated stationary distribution of workers Γ_k solves $\Gamma_k = \boldsymbol{Q}_k \circ \Gamma_k$.

C Figures and tables omitted from the main text



(a) Self-employment rate



(b) Unemployment rate

Figure C.1: Benchmark SIPP vs Bureau of Labor Statistics Series (BLS). The self-employment and unemployment rate in the SIPP are based on the definitions of labor force status given in section 2 and appendix A. The sample used to build the series include all working age individuals, not just the main earner in each household. Observations are aggregated with cross-sectional weights.

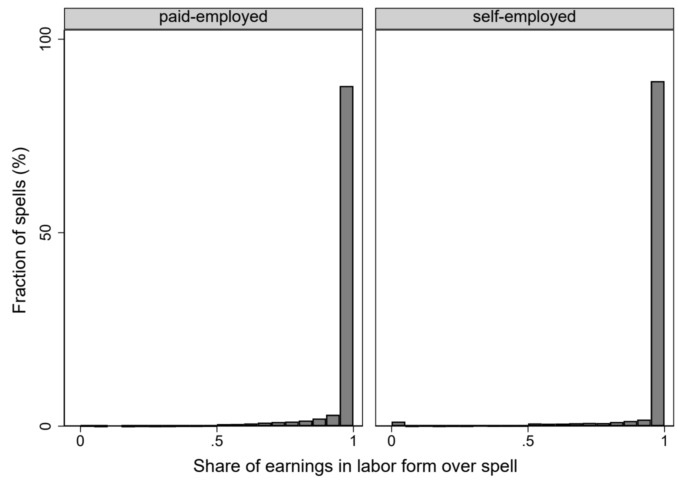


Figure C.2: Share of labor earnings over employment spell, conditional on ascribed labor form.

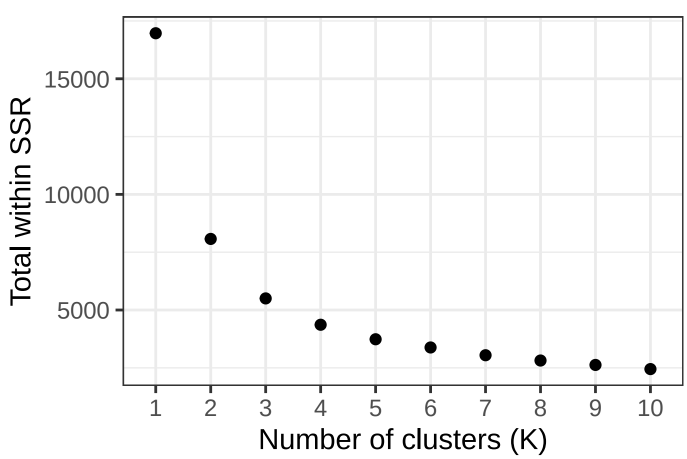


Figure C.3: Change in kmeans fit with number of clusters. The measure of fit is the total within sum of square residuals (SSR).

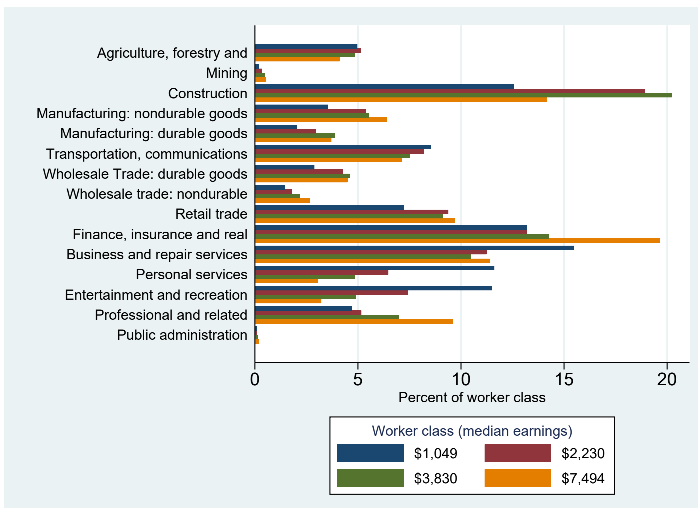


Figure C.4: Industry composition by worker class.

Reported business wealth (Median, \$2009)	
Assets	6,355
Debt	0
Equity	4,770
Share incorporated	0.320
Share with more than 25 employees	0.050

Table C.1: Business characteristics. Note: Business wealth is obtained by multiplying the owner's share and summing the data for all businesses, if several businesses are reported. The dummy variables take value one if any of the reported businesses satisfies the condition.

	Worker class k (median earnings)			
$\ln Y_{it}^s = \alpha_k^s + \beta_k^s \cdot \ln y_{it}^s + \epsilon_{it}$	\$1,032	\$2,227	\$3,840	\$7,626
Unemployment ($s = U$)				
α_k^U	6.720	6.849	6.892	7.209
Paid-employment ($s = P$)				
α_k^P	3.460	2.978	2.894	2.437
β_k^P	0.601	0.677	0.699	0.762
Self-employment ($s = S$)				
α_k^S	3.277	2.906	2.858	2.678
β_k^S	0.629	0.688	0.705	0.736

Table C.2: Household income functions.

Parameter	Description	Worker class k (Median earnings in \$2009)			
		1,032	2,227	3,840	7,626
δ_k^P	Destruction rate in P	0.041	0.028	0.018	0.015
δ_k^S	Destruction rate in S	0.009	0.006	0.004	0.002
μ_k^P	Income draw in P:	7.022	7.647	8.206	8.864
σ_k^P	$\ln \mathcal{N}(\mu_k^P, \sigma_k^P)$	0.232	0.405	0.327	0.392
μ_k^S	Income draw in S:	6.778	7.656	8.221	8.936
σ_k^S	$\ln \mathcal{N}(\mu_k^S, \sigma_k^S)$	0.726	0.574	0.563	0.497

Table C.3: Parameters—Specific to worker classes.

Policy	Benefits	Contributions	Equilibrium τ
Unemployment insurance (UI)			
UI-A (align)	$\max\{.5 \cdot y, 2000\}$ in $s = C$	$\min\{\tau \cdot y, 7000/12\},\ s = S$	0.025
UI-P (progressive)	$\max\{.5 \cdot y, 2000\}$ in $s = C$	$\tau \cdot y,\ s = S$	0.003
Basic income (BI)	$b = 100,\ \forall s$	$\tau \cdot Y^s,\ \forall s$	0.014

Table C.4: Summary of policy experiments.