Labor Income Shocks along the Business Cycle

Click here for the latest version

Julien Pascal *†

January 20, 2020

Abstract

This paper analyzes the determinants of labor income shocks along the business cycle. My main finding is that sorting between firms and workers is a key component of idiosyncratic risk. Labor income shocks are analyzed through the lenses of a dynamic search-and-matching model, which I estimate using US data. Because of search frictions and mismatches between firms and workers, the *laissez-faire* equilibrium is not necessarily optimal. My results underline that the government can tame business cycle fluctuations by designing a simple unemployment policy improving sorting between firms and workers.

Keywords: Sorting, Labor Income Risk, Business Cycle

JEL Classification: E32, J31

*I would like to thank Jean-Marc Robin, Xavier Ragot, Florian Oswald, Pierre Cahuc, Wouter Den Haan, Jeremy Lise and Nicolò Dalvit and all my colleagues at Science Po for valuable comments and feedback. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program within the framework of the LIEPP center of excellence (ANR11LABX0091, ANR 11 IDEX000502)

†Julien Pascal, Sciences Po, Department of Economics, LIEPP, Paris, France

1 Introduction

This paper has a triple objective. The first one is to deepen our understanding on the sources of labor income shocks. That is, the unpredictable part of labor income changes. It is now well established that fluctuations in earnings at the individual level is an order of magnitude bigger than fluctuations at the macro level (Parker and Vissing-Jorgensen (2009)). The assumption of normality of labor income shocks has been attacked by several recent publications, in particular in Guvenen et al. (2014) and Guvenen et al. (2015). Recessions are periods marked by intense negative labor income shocks, underlined by a spike in left-skewness in the distribution of labor income changes. The aim of this paper is to unpack the black-box of the complex labor income process and to analyze its determinants. Why should we care? What does left-skeweness mean for an average worker? In practical terms, it means that some categories of workers are badly hurt by recessions, with persistent consequences. The scarring effects of recessions are now well identified (see Ouyang (2009)). The persistence of labor income shocks can be in part explained by search models with human capital depreciation, which creates a negative feedback loop on aggregate variables, as in Walentin and Westermark (2018). If negative feedback loops are involved, preventing bad shocks from happening or helping workers to recover from them is probably a welfare enhancing policy.

This paper sheds light on a previously ignored component of idiosyncratic income risk: sorting between workers and firms. By sorting, I mean the extent to which the market allocates the right workers to the right jobs, where "right" is captured by complementarities in the production function. Why is sorting an important mechanism for the labor income process? When considering a labor market with search frictions and random search, the pairing between firms and workers is not necessarily optimal. An inefficient match in turn translates into lower wages as long as the match persists. Being fired not only has a direct consequence on someone's labor income, it also has dynamic consequences. To go back to her/his previous income level, a newly unemployed worker has to climb up the *intra-firm wage ladder* and the *inter-firm ladder*. That is, a worker has to spend some time searching on the labor market before finding a firm that is the right match, and conversely. Numerical simulations show that the inter-firm ladder is far from being negligible. Long-tails in the distribution of labor income shocks hinges on the economy featuring heterogeneous firms, hence an inter-firm ladder. To the best of my knowledge, empirical work on this component of risk is rather scarce. One notable exception is Huckfeldt et al. (2016), who shows that earnings costs of job loss are concentrated among workers who find reemployment in lower-paying occupations. Using CPS and PSID data, the author estimates that the initial earnings losses of workers losing their job and subsequently switching occupations are *four times larger* than losses for workers finding a job in their previous occupation. Persistence of the initial wage loss is only observed for occupation switchers. These empirical facts can be consistently explained by the existence of an inter-firm ladder combined with some degree of random search on the labor market.

Related literature on sorting includes the seminal contribution of Abowd et al. (1999), decomposing real total annual compensation per worker into an employee, an employer and a residual effect. Bonhomme et al. (2019) introduce a framework that can accommodate interactions between worker and firm attributes. In a variance decomposition exercise, Song et al. (2018) show that two-thirds of the rise in the dispersion of log earnings between 1978 and 2013 can be attributed to a rise in the dispersion of average earnings *between* firms. In the search-and-matching literature, Lise and Robin (2017) study how sorting patterns are altered along the business cycle. To study labor income shocks across the cycle, my strategy is to use and extend their model. My contribution is twofold. Firstly, I extend the model of Lise and Robin (2017) by solving for the wages, which were left implicit in their contribution. Secondly, I estimate the model using the simulated method of moments (SMM) and wage moments. In particular, I focus on starting wage moments, which are particularly well-defined within the model and have strong identification power.

A second objective of this paper explores new techniques to solve and estimate dynamic search-and-matching models with heterogeneity. Solving labor models with both search frictions and heterogeneous agents is a notoriously difficult task. If workers do not have access to full information on the state of the labor market, which includes the number of vacancies posted by each firm and their associated wage, a commonly held view is that such frameworks cannot be solved using standard numerical techniques. To avoid these complications, the literature on dynamic search-and-matching models has focused on *block-recursive* equilibrium, following the seminal contribution of Menzio and Shi (2010). In short, a block-recursive model is one in which value functions and market tightness are independent from the distribution of employment across worker types. Such knife-edge conditions are met when search is *directed*. That is, (i) when firms make public the wage associated to each vacancy they post (ii) workers have full information over wages and the types of vacancy posted. Armed with full knowledge of the labor market conditions, workers direct their search efforts towards a specific sub-market.[1] While particularly clever and numerically efficient, block-recursive models are constrained efficient as a by-product of the modeling tricks involved (see Schaal (2017)). Thus, in a block recursive model, the *laissez-faire* equilibrium is necessarily optimal. When the goal of a paper is to explain a mechanism through the eyes of a model, constrained efficiency is mostly harmless. However, if the objective is to understand how a government may or may not improve the market outcome, it seems more appropriate to come up with a new concept of equilibrium that does not rule out inefficiencies in the first place. This is the route I explore in this paper. The strategy I use to solve a non block-recursive search model can be seen as a variant of the Krusell and Smith (1998) algorithm. Agents are endowed with a simple forecasting rule that needs to be estimated via Monte-Carlo. The particularity in my setting is that the time series needed to estimate the forecasting rule do not depend on the value functions to be calculated. This property, specific to the model under scrutiny, leads me to design an algorithm that is both rapid and robust to solve the model. Because the model is half way between a fully non-block recursive model and a block recursive one, I propose to name it as *semi-block recursive*.

To a lesser extent, my work is related to the burgeoning literature on how to solve and estimate models with both aggregate uncertainty and heterogeneity. Following the seminal contribution of Krusell and Smith (1998), several techniques have been developed (see Reiter (2009), Algan et al. (2014) and Winberry (2018)). While some of these techniques are global and the other ones use linearization around the non-stochastic steady-state, the common denominator of the above mentioned methods is that they rely on the recursive representation of a multi-stage decision process. An interesting line of research has recently used the sequence representation of the dynamic choice problem (see Le Grand et al. (2017), Boppart et al. (2018) and Auclert et al. (2019)). In this paper, I stick to the more commonly used recursive form, but I note that the sequence form is particularly

[1] A free entry condition for firms is also needed.

well-suited in my setting. My contribution is not to develop a new method per se, but to realize that there is a space in between models that are fully non-recursive and models that are fully recursive. Within that thin space, exiting methods can easily be applied.

A third objective of this paper is to analyze the potential gains from designing an optimal unemployment insurance (UI). Having defined a concept of equilibrium in which the *laissez-faire* equilibrium is not optimal by design, I explore simple unemployment policies that have the potential of being welfare-improving. The design of an optimal UI and the extent to which it can stabilize the cycle has received a comprehensive treatment in the macroeconomic literature. In a framework with heterogeneous agents and aggregate uncertainty, Ragot and Le Grand (2019) solve for the optimal Ramsey problem. The optimal replacement rate is pro-cyclical and stabilizes aggregate demand. In the present paper welfare gains are realized by improving the improving the sorting between firms and workers, boosting the value of production. In a similar setting, Lise et al. (2016) show how an optimal replacement raste might improve the market equilibrium and transfer utility across groups of workers. They find that the optimal unemployment scheme can deliver a welfare improvement of 1.4%, concentrated on low-skill workers. I contribute to this literature by analyzing a similar unemployment insurance scheme in a dynamic setting. I find that the optimal unemployment insurance scheme generates a 0.25% increase in welfare at the steady-state. While the gains are rather modest at the steady-state, the policy is successful in stabilizing labor income shocks over the business cycle by approximately 2%. These gains are achieved by transferring income from high-skilled to low-skilled workers and by a stabilization of the inter-firm channel. The mechanism is quite intuitive: by making unemployment workers better off, especially low-skilled workers, they become more selective when choosing a job. The congestion effects of low-skilled workers are mitigated and high-skilled workers end up in better matches. As high-skilled workers are less likely to lose their job when the economy enters a recession, fluctuations in labor income are less severe.

2 Data

In this section, I calculate statistics on the wage distribution along the business cycle. I focus on the elasticity and the standard deviation of wages for the entire workforce and for the workers who just exited unemployment. Putting the spotlight on the wage of newly employed workers is important for both theoretical reasons linked to the Shimer's puzzle (see Shimer (2005) and Pissarides (2009)) and for practical reasons linked to my modeling strategy. In the model developed below, the distribution of wages for new hires is well-behaved and has strong identification power. I first describe the methodology used to calculate starting wage moments, which closely follows Haefke et al. (2013), before analyzing the determinants of labor income shocks in the US.

2.1 Wage dynamics

To calculate statistics on the starting wage distribution in the US, I use the CPS Merged Outgoing Rotation Groups (CPS MORG), which contain both employment and wage variables for the period 1979 until nowadays.[2] More specifically, I use the Center for Economic and Policy Research ORG extracts[3], which contain time-consistent CPS MORG variables and wage variables corrected for top-coding in declared wages. To measure real hourly wage, I use the CEPR MORG variable *rw*, which excludes overtime, tips and commissions for hourly workers; but includes overtime, tips and commissions for non-hourly workers. The dataset is trimmed to excludes observations where real 1989 hourly wage is smaller than \$0.50 or bigger than \$200.

While the CPS was designed to offer repeated cross-section views of the US population, it also has a longitudinal dimension. Every household that enters the CPS is interviewed each month for 4 months, then ignored for 8 months, then interviewed again for 4 more months. Usual weekly hours/earning questions are asked only at households in their 4th and 8th interviews.[4] Hence, by comparing the same individual's hours/earnings in the 4th and 8th interviews, one can calculate the evolution of hourly wage over a year period. One can also determine which workers transitioned from unemployment to employment in between two interviews.

The process is complicated by the fact that keeping track of individuals in the CPS MORG is not straightforward, as the unit of reference is a housing unit. About 60,000 housing units are designated for data collection each month. Each house is assigned

[2]Respondents are asked to report their hourly wage if they are hourly workers or to report their weekly earning if they are paid by the hour.

[3]http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-programs/

[4]See https://www.nber.org/data/morg.html

a household identifier (HHID) and an individual within the household is assigned an individual line number (LINENO). Individuals not changing location in between two interviews are in theory uniquely identified by the pair (HHID-LINENO). When a new household moves in, a "household counter" variable (HHNUM) is incremented by 1. To control for households moving in, I build an individual identifier as the combination of (HHID-LINENO-HHNUM) and immutable characteristics (gender and ethnicity). As an additional safety check, I compare the age of each potential individual between two observations. If the age difference is less than or equal to 2, I validate the match. Otherwise, the match is discarded. By choosing an age difference of 2 instead of 1, I allow for some degree of coding error. Using the (HHID-LINENO-HHNUM-gender-ethnicity) identifier, I can keep track of changes in the labor force status for individuals (employed, unemployed or not in the labor force). I calculate deciles at the yearly frequency for the starting wage distribution (workers currently employed and previously unemployed or out of the labor force) and for the entire distribution of wages.

Key statistics to estimate the models are the volatility and elasticity of (real) wage deciles with respect to changes in labor productivity. The response of wage deciles to productivity is measured by the coefficient of a regression of the log real wage deciles on log real labor productivity. As in Haefke et al. (2013), I estimate the regression in first differences to avoid spurious correlation if wages and productivity are integrated:

$$\Delta \log(w_{d,t}) = \alpha_d + \eta_j \Delta \log(y_t) + \varepsilon_{d,t} \tag{1}$$

where $w_{d,t}$ denotes the dth wage decile at time t and y_t is a measurement of labor market productivity at time t.[5] I estimate equation (1) on a sample restricted to men and women in between 25 and 60 years old working in the private sector, with series aggregated at the yearly frequency. While data is available starting in 1979, I restrict the sample to a period starting with the Great Moderation in 1984, as in Haefke et al. (2013). The period 1979-1983 is marked by large volatility in macro variables and by a substantial drop in the real minimum wage, which pushes downward the estimates of Table 1. Results are presented in Table 1. Consistent with the empirical literature (see Pissarides (2009)), I find that wages for new hires are much more sensitive to variations in

[5] To measure labor productivity, I use real output per hour of all persons in the non-farm (OPHNFB)

labor productivity than in the series for aggregate wages. My point estimates are slightly less than reported by Haefke et al. (2013)[6], but of the same order of magnitude. Wage rigidity in on-going contracts is important because it speaks against using a continuously-renegotiated Nash bargaining to determine wages. If wages are the solution of the Nash sharing rule, every movement in aggregate productivity leads to variations in workers' wages, which is at odds with the data.

To estimate the volatility of each wage decile, I first detrend series using a linear trend or a HP filter. Results are presented Table 2. Two facts are worth noting. Firstly, wage deciles for new hires are much more volatile than the overall population. Depending on whether detrending is done using a linear trend or an HP filter, the volatility of the median wage for new hires is between 40% and 100% higher than the general population. Secondly, when considering the sample of all workers, lower percentiles tend to be more volatile compared to the top of the wage distribution. This pattern is easily explained by the fact that the lower percentiles of the wage distribution are predominantly impacted by the wage of new hires, which are more volatile than the wage of workers in ongoing contracts.

[6]When the authors do not control for education (as it is the case in this paper), they find an aggregate wage elasticity of 0.16 for all workers and 0.54 for new hires for the period 1984 - 2006. When the authors control for difference in obervables characteristics, they find and elasticity of 0.24 for all workers and 0.79 for new hires.

Table 1: Elasticity of wage and starting wage deciles in the US: 1984 - 2017

	New hires		All workers	
	η_j	p-value	η_j	p-value
P10	0.44	0.07	0.20	0.37
P20	0.32	0.27	0.13	0.45
P30	0.34	0.21	0.18	0.29
P40	0.31	0.32	0.18	0.24
P50	0.35	0.21	0.12	0.41
P60	0.42	0.09	0.30	0.06
P70	0.37	0.17	0.26	0.04
P80	0.45	0.20	0.28	0.02
P90	0.65	0.08	0.46	0.00

Notes: This table shows p-values and point estimates for the regression (1), measuring the sensitivity of wage deciles to changes in aggregate labor productivity. To measure productivity, I use real output per hour of all persons in the non-farm sector (OPHNFB). To measure hourly wage, I use the series 'rw' from the CEPR CPS ORG Extract, which converts hourly pay to constant 2018 dollars using the CPI-U-RS and corrects for top-coding. For calculations involving starting wages, the years 1986, 1995 and 1996 are excluded from the sample. For years 1986 and 1995-1996, I find only a limited number of workers transiting from unemployment to employment relative to other years (less than 1250 workers). Digits were rounded to the nearest hundredth.
Sources: CEPR CPS ORG Extract (http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/) and U.S. Bureau of Labor Statistics retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/OPHNFB

Table 2: Volatility of wage and starting wage deciles in the US: 1984 - 2017

	New Hires		All workers		Volatility Ratio New hires/All workers	
	Linear trend	HP-filter	Linear trend	HP-filter	Linear trend	HP-filter
P10	0.041	0.019	0.036	0.017	1.14	1.14
P20	0.044	0.021	0.038	0.011	1.17	1.98
P30	0.051	0.019	0.029	0.012	1.77	1.61
P40	0.049	0.024	0.028	0.010	1.75	2.27
P50	0.039	0.020	0.028	0.010	1.39	1.98
P60	0.043	0.018	0.026	0.012	1.65	1.53
P70	0.041	0.020	0.022	0.010	1.85	2.00
P80	0.047	0.025	0.022	0.009	2.11	2.81
P90	0.045	0.026	0.024	0.010	1.82	2.59

Notes: This table the standard deviation of deciles of the log real wage distribution for new hires and all workers. The log of wage deciles were detrended using a linear trend or using a HP-filter with a smoothing parameter equal to 6.5. For calculations involving starting wages, the years 1986, 1995 and 1996 are excluded from the sample. For these years, I am able to find only a limited number of workers transiting from unemployment to employment relative to other years (less than 1250 workers).
Sources: CEPR CPS ORG Extract (http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/).

2.2 Labor Income Shocks

Labor income shocks are defined as the unpredictable part of labor income changes. Let $w_{i,t}$ denote the real hourly wage of individual i at time t. I first project $w_{i,t}$ on a set of observable characteristics:

$$\log(w_{i,t}) = \boldsymbol{x_{i,t}}'\boldsymbol{\beta_w} + \varepsilon_{i,t} \tag{2}$$

where $x_{i,t}$ includes a constant, age and age square, marital status, education level and a linear time trend to capture long-run dynamics impacting real wage. Results are presented in Table 3. Log hourly wage is an increasing and concave function of age; a higher education level is associated with a higher hourly wage; holding other factors constant, women earn 28% less than men. By construction, the residual $e_{i,t} \equiv \log(w_{i,t}) - \boldsymbol{x_{i,t}}'\boldsymbol{\hat{\beta}_w}$ is orthogonal to the set of observable variables included in the right hand side of equation (2). I construct an hourly wage index orthogonal to observable characteristics as follows:

$$\log(\tilde{w}_{i,t}) = \log(w_{i,t}) + (\boldsymbol{\bar{x}'} - \boldsymbol{x_{i,t}}')\boldsymbol{\hat{\beta}_w} \tag{3}$$

where $\boldsymbol{\bar{x}'}$ denotes the average observable characteristics across individuals and periods. By construction $\Delta \log(\tilde{w}_{i,t}) \equiv \log(\tilde{w}_{i,t}) - \log(\tilde{w}_{i,t-1})$ measures the (log) difference in hourly wage that cannot be explained by observable factors. Note that by taking the difference for the same individual i, unobservable individual fixed effects that may have explained parts of labor income changes are removed. I apply the same methodology for weekly hours worked, by first fitting a linear model and then removing predictable changes in hours. Point estimates for the linear model are presented in table 4. Let $\Delta\tilde{h}_{i,t}$ denote the unpredictable part of changes in weekly hours worked, which is calculated according to the following formula:

$$\Delta\tilde{h}_{i,t} = (h_{i,t} - h_{i,t-1}) - (\boldsymbol{x_{i,t-1}}' - \boldsymbol{x_{i,t}}')\boldsymbol{\hat{\beta}_h} \tag{4}$$

I generate a measurement of weekly labor income orthogonal to changes in observable factors, denoted by $\tilde{y}_{i,t}$, using $\tilde{w}_{i,t}$ and $\tilde{h}_{i,t}$:

$$\tilde{y}_{i,t} = \tilde{h}_{i,t} \times \tilde{w}_{i,t} \tag{5}$$

The resulting $\tilde{y}_{i,t}$, as well as $\tilde{h}_{i,t}$ and $\tilde{w}_{i,t}$ and are reported in Figure 5. Visual inspection of Figure 5 indicates that recessions years are marked by a contemporaneous drop in mean hours worked. The early 1980s and 1990s recessions were characterized by a contemporaneous drop in the mean real hourly wage, but the early 2000s recession and the Great Recession of 2008-2009 were inflexion points, with a decrease in the mean real hourly wage following with a lag. To complement this picture at the aggregate level, I calculate the probability that a worker experiences certain events. At the individual level, the probability of losing more than 0.5 times the standard deviation in real weekly earnings (approximately \$579) jumps by 1.24 percentage points in recession (see tables 5 and 6). The drop in real weekly earnings is caused by a decrease in hourly wage and a decrease in hours worked. Large negative changes in hours worked are more frequent in recession[7] and the probability of a large increase in hours worked declines.

[7] A 2σ decrease in hours worked corresponds to a drop in weekly hours worked of approximately 19 hours.

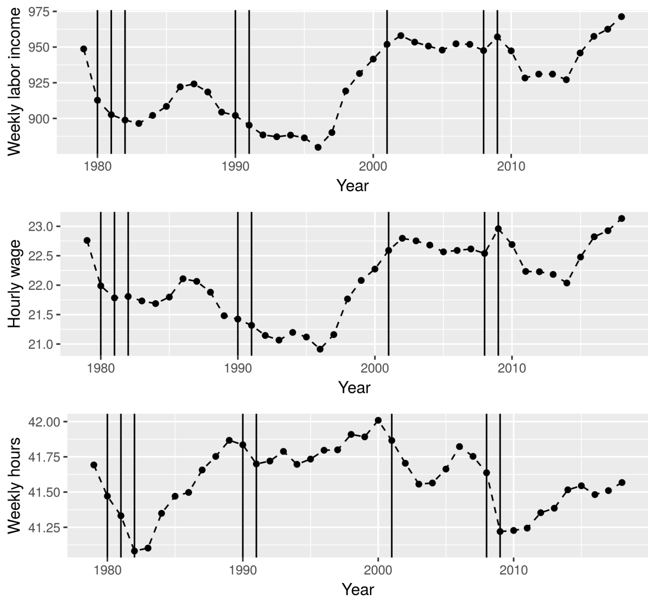
Table 3: Regression log wages

	Dependent variable: log hourly wage $\log(w_{i,t})$
age	0.044***
	(0.0002)
age^2	−0.0004***
	(0.00000)
married	0.080***
	(0.0005)
trend	0.001***
	(0.00002)
HS	0.242***
	(0.001)
some college	0.392***
	(0.001)
college	0.691***
	(0.001)
advanced	0.868***
	(0.001)
woman	−0.280***
	(0.0004)
constant	1.601***
	(0.004)
Observations	5,309,050
R^2	0.281
Adjusted R^2	0.281
Residual Std. Error	0.487 (df = 5309040)
F Statistic	230,227.100*** (df = 9; 5309040)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression weekly hours worked

	Dependent variable: weekly hours worked $h_{i,t}$
age	0.280***
	(0.004)
age^2	−0.003***
	(0.00005)
married	−0.438***
	(0.009)
trend	−0.004***
	(0.0004)
HS	0.776***
	(0.015)
some college	1.101***
	(0.016)
college	2.495***
	(0.017)
advanced	4.315***
	(0.019)
woman	−5.398***
	(0.008)
constant	35.462***
	(0.077)
Observations	4,735,952
R^2	0.096
Adjusted R^2	0.096
Residual Std. Error	9.087 (df = 4735942)
F Statistic	55,871.350*** (df = 9; 4735942)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 1: Weekly earnings, hourly wage and hours worked



Notes: This figure shows the year-specific mean values for the residualized weekly hours worked $\tilde{h}_{i,t}$, the residualized real hourly wage $\tilde{w}_{i,t}$, and the implied weekly labor income $\tilde{y}_{i,t} = \tilde{h}_{i,t} \times \tilde{w}_{i,t}$. Vertical lines represent NBER recessions.

Table 5: Probability of changes in hourly wage, weekly earnings and hours along the business cycle

	$\tilde{w}_{i,t}$	$\tilde{w}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{h}_{i,t}$	$\tilde{h}_{i,t}$
Recession	No	Yes	No	Yes	No	Yes
$\Pr(\Delta x_{it} < -2\sigma)$	2.36%	2.37%	2.63%	2.67%	3.81%	4.00%
$\Pr(\Delta x_{it} < \sigma)$	9.20%	9.64%	9.77%	10.45%	10.82%	10.94%
$\Pr(\Delta x_{it} < 0.5\sigma)$	19.47%	20.26%	19.99%	21.23%	18.14%	18.78%
$\Pr(\Delta x_{it} > 0.5\sigma)$	20.56%	20.37%	21.24%	21.04%	18.84%	17.72%
$\Pr(\Delta x_{it} > \sigma)$	9.68%	9.72%	10.36%	10.34%	11.37%	10.40%
$\Pr(\Delta x_{it} > 2\sigma)$	2.54%	2.42%	2.85%	2.66%	4.04%	3.90%

Notes: This table displays the probability that a yearly difference in x_{it} is above or below certain threshold σ, which denotes the standard deviation of Δx_{it}. $\tilde{w}_{i,t}$ denotes the hourly wage of worker i at time t, net of net of predictable factors (age, education, etc). $\tilde{h}_{i,t}$ denotes weekly hours worked net of predictable factors. $\tilde{y}_{i,t}$ denotes the weekly labor income of individual i at time t.

Table 6: Difference in the probability of changes in hourly wage, weekly earnings and hours

	$\tilde{w}_{i,t}$	$\tilde{y}_{i,t}$	$\tilde{h}_{i,t}$
$\Delta\,\Pr(\Delta x_{it} < -2\sigma)$	0.01%	0.04%	0.19%
$\Delta\,\Pr(\Delta x_{it} < -\sigma)$	0.44%	0.68%	0.12%
$\Delta\,\Pr(\Delta x_{it} < -0.5\sigma)$	0.78%	1.24%	0.64%
$\Delta\,\Pr(\Delta x_{it} > 0.5\sigma)$	-0.19%	-0.20%	-1.12%
$\Delta\,\Pr(\Delta x_{it} > \sigma)$	0.03%	-0.02%	-0.97%
$\Delta\,\Pr(\Delta x_{it} > 2\sigma)$	0.13%	-0.19%	-0.14%

Notes: This table displays the difference (recession minus expansion) in the probability that a yearly difference in x_{it} is above or below certain threshold. The threshold value σ denotes the standard deviation of Δx_{it}. $\tilde{w}_{i,t}$ denotes the hourly wage of worker i at time t, net of net of predictable factors (age, education, etc). $\tilde{h}_{i,t}$ denotes weekly hours worked net of predictable factors. $\tilde{y}_{i,t}$ denotes the weekly labor income of individual i at time t.

3 Model

The empirical section underlined that (i) wages of newly hired workers are correlated with productivity, while wages in ongoing contracts are quite rigid (ii) short-term downside risks in labor income are mainly driven by losses in hours. In this section, I develop a model that delivers both features and that allows me to decompose labor income risk between a part that is driven by workers' characteristics and another that depends on firms. To model the employment side, I choose the framework of Lise and Robin (2017), which comes with a natural notion of sorting. Because search is random, workers may not necessarily meet with their optimal firm type, as measured by complementarities in the production function. An alternative framework featuring two-sided heterogeneity and aggregate uncertainty is the model of Schaal (2017), built upon the directed search model of Menzio and Shi (2010). However, it is hard to define sorting in the latter framework because the directed search assumption generates a constrained efficient market outcome. Conditionally on the state of the economy, workers are always is the best match the can achieve.

My contribution in terms of modeling is to develop an efficient technique to solve for the wages. The key insight is that (i) the employment problem is independent from determination of wages (ii) while the wage process depends on the employment process, this dependency is rather mild. That is, while the wage problem is not recursive (the state variable is infinite dimensional), a dimension reduction in the spirit of Krusell and Smith

(1998) can be used. The wage problem is particularly well-behaved because the time series needed for the dimension-reduction step are independent from the value function for wages.

3.1 Workers, Firms and Timing

The economy is populated by a continuum of risk-neutral and infinitely-lived workers with mass 1. A worker can be either employed or unemployed. Workers differ in a skill parameter x distributed according to a density $\ell(x)$. There is a continuum of risk-neutral firms, differing by a productivity parameter y uniformly distributed on $[0, 1]$. Firms do not need capital to operate and can only hire one worker at a time. Firms can freely enter the market and do so until the value of an unfilled vacancy is zero. Firms advertise positions through job placement agencies. The cost of posting v vacancies is given by a strictly increasing and convex cost function $c(.)$. In equilibrium, the marginal cost of creating a vacancy is equal to the expected return of doing so. Aggregate uncertainty stems from an aggregate productivity parameter z that follows an $AR(1)$ process:

$$z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{t+1} \tag{6}$$

with ε_{t+1} an i.i.d. Gaussian random variable with zero mean and unit variance. The value of home production for a worker of type x is given by $b(x, z_t)$ and the value of output for a worker of x working with a firm of type y is denoted by $p(x, y, z_t)$. The timing is as follows. The measure of $x - y$ matches (jobs) at the beginning of period t is denoted by $h_t(x, y)$. Because workers are either employed or unemployed, the following accounting identity holds:

$$\ell(x) = u_t(x) + \int_0^1 h_t(x, y)dy \tag{7}$$

where $u_t(x)$ denotes the measure of workers of type x unemployed at the beginning of period t. The aggregate productivity variable changes from z_{t-1} to z_t. Right after the change in productivity, workers may lose their job for exogenous or endogenous reasons. Job search, matching and wage setting happen in a sub-period $t+$. The measure of matches surviving job destruction is denoted by $h_{t+}(x, y)$ and the measure of unemployed workers in the sub-period $t+$ is denoted by $u_{t+}(x)$. Unemployed workers and employed workers both produce a search effort L_t, which is linear aggregation of individual efforts:

$$L_t = \int_0^1 u_{t+} dx + s \int_0^1 \int_0^1 h_{t+}(x, y) dx dy \tag{8}$$

The free entry condition on the firm's side implies that the marginal cost of posting one vacancy is equal to the expected value of a job opening:

$$c'(v_t(y)) = q_t J_t(y) \tag{9}$$

where $J_t(y)$ denotes the expected value of a contact by a vacancy of type y and q_t is the probability (per recruiting effort) that a firm contacts a worker. Because the cost function is assumed to be increasing and convex, $c'(.)$ can be inverted $v_t(y) = (c')^{-1}(q_t J_t(y))$. The total number of vacancies in period t is obtained by integrating over firm types $V_t = \int v_t(y) dy$. The total number of meetings at time t, denoted by M_t, is the result of workers' search efforts and firms' vacancy posting behavior. A matching function $M(.)$ is used to model the meeting of both sides of the labor market:

$$M_t = M(L_t, V_t) \tag{10}$$

The probability for a worker to meet a vacancy is the ratio of the number of meeting to the aggregate search effort $\lambda_t = \frac{Mt}{L_t}$. The probability that a firm contacts any searching work q_t is the ratio of the number of meeting to the total number of vacancies $q_t = \frac{Mt}{V_t}$.

3.2 Wage setting

Wages are determined according to the sequential auction framework, as in Robin (2011) or Postel-Vinay and Turon (2010). Unemployed workers have zero bargaining power and receive their reservation wage. Employed workers search for alternative employers. When they meet another firm, workers reveal the meeting to their current employer. A Bertrand competition between the incumbent and the poaching firm is triggered, which results in either a wage increase and/or a job-to-job transition.

In this environment, two properties are absolutely essential. Firstly, the value of unemployment to a worker of type x when the aggregate productivity level is z_t, denoted by

$U(x,z_t)$, is independent from the distribution of matches $h_t(x,y)$. Secondly, the joint surplus of a match, denoted by $S(x,y,z)$, does not depend on $h_t(x,y)$ either. The functions $U(.)$ and $S(.)$ are solution of the following functional equations:

$$U(x,z_t) = b(x,z_t) + \frac{1}{1+r}\,\mathbb{E}_{z_{t+1}|z_t}\left[U(x,z_{t+1})\right] \tag{11}$$

$$S(x,y,z_t) = p(x,y,z_t) - b(x,z_t) + \frac{1-\delta}{1+r}\,\mathbb{E}_{z_{t+1}|z_t}\left[\max\left(0, S(x,y,z_{t+1})\right)\right] \tag{12}$$

where r is the interest rate. The expectation operator is taken with respect to next period's aggregate productivity level only. Equations (11) and (12) can be trivially solved by value function iteration.

Two points are worth emphasizing. Firstly, the joint surplus of match $S(x,y,z_t)$ does not depend on the job meeting rate λ_t. Independence from the job meeting rate in turn implies the joint surplus of a match does not depends on the distribution of matches across skill and firm productivity types $h_t(x,y)$, an infinite dimensional object. Independence of $S(x,y,z_t)$ from the job meeting rate hinges on unemployed workers having zero bargaining power. Secondly, the joint surplus of a match does not depend on wages. The fact that $S(x,y,z_t)$ does not depend on wages rests on the assumption that workers and firms are risk neutral. Within a match, the wage is an instrument to decide the split of the joint surplus between workers and firms. Because utility is linear, the allocation of the surplus between the two parties does not modify the surplus itself. Independence of the joint surplus of a match from $h_t(x,y)$ and $h_t(x,y,w)$ are both essential when developing an algorithm to solve efficiently the wage process.

Knowledge of the joint surplus of a match $S(x,y,z_t)$ is sufficient to determine both job feasibility and job-to-job movements. Using this fact, the employment side of the model can be closed. The assumption of zero bargaining power for unemployed workers and the sequential auction hypothesis yield the following expression for the the expected value of a contact:

$$\begin{aligned} J_t(y) = &\int_0^1 \frac{u_{t+}(x)}{Lt} \max\left(0, S(x,y,z_{t+1}\right) dx \\ &+ s\int_0^1\int_0^1 \frac{h_{t+}(x,y')}{L_t}\max\Big(0, S(x,y,z) - S(x,y',z)\Big)dxdy \end{aligned} \tag{13}$$

The first term is the expected value of hiring from the pool of unemployed workers, while the second term is the expected value of poaching workers from less productive firms. The measure of $x - y$ matches in the sub-period $t+$, surviving exogenous and endogenous job destruction is:

$$h_{t+}(x,y) = (1-\delta)\mathbb{1}\{S(x,y,z_t) \geq 0\}h_t(x,y) \tag{14}$$

The measure of matches at the end of period t takes into account the measure of workers moving to more productive poachers, the measure of workers poached from less productive firms, and the inflow of workers hired from unemployment:

$$\begin{aligned} h_{t+1}(x,y) = h_{t+}(x,y)\Big[1 - s\int_0^1 \lambda_t \frac{v_t(y')}{V_t}\mathbb{1}\{S(x,y',z_t) > S(x,y,z_t)\}dy'\Big] \\ + s\int_0^1 h_{t+}(x,y')\lambda_t \frac{v_t(y)}{Vt}\mathbb{1}\{S(x,y,z_t) > S(x,y',z_t)\}dy' \\ + u_{t+}(x)\lambda_t \frac{v_t(y)}{Vt}\mathbb{1}\{S(x,y,z_t) \geq 0\} \end{aligned} \tag{15}$$

4 Solving for the wages

Let $W_t(x,y)$ denote the value of a job to a worker x working with a firm of type y at time t. In the class of models with search frictions and aggregate uncertainty, $W_t(x,y)$ generally depends on next period's job meeting rate λ_{t+1}. If the job meeting rate is high, an employed worker is more likely to receive a promotion or to change job. The job meeting in the future thus impacts the reservation wage of workers today. The job meeting rate is itself a function of the current distribution of matches $h_t(x,y)$. Indeed, $h_t(x,y)$ affects both firms' expected value of posting vacancies (equation (13)) and workers' search effort (equation (8)). If no additional assumptions are made, the relevant aggregate state variable for the determination of wages contains *the joint distribution of matches* $h_t(x,y)$.

A convenient assumption is to posit that a contract is an agreement to receive a given share of the match surplus $S(x,y,z_t)$. Given that the surplus does not depend on the distribution of matches $h(x,y)$, this property is inherited by $W_t(x,y)$. This path is followed by Lise et al. (2017). Yet, this assumption implies that the wage changes every time z does, even when workers or firms have no credible threat to quit or change job. The empirical analysis underlined that wages in ongoing contracts are rigid, which suggests that an alternative route could be considered. One alternative is to define a contract as an agreement to a constant wage w, which can be re-bargained by mutual consent only. This is the path I explore in this paper. I resolve the difficulty of having $h_t(x,y)$ in the aggregate state variable by realizing that the dependency of $W_t(x,y)$ on $h_t(x,y)$ is rather mild. As a result, dimension reduction tools from the literature on heterogeneous agents can be easily applied.

4.1 The value of a job to workers

The value of a y job to a worker of type x, with a wage equal to w, when the aggregate state variable is $\Gamma_t \equiv \big(z_t, h_t(x,y)\big)$ writes:

$$\begin{aligned} W(x,y,w,\Gamma_t) = u(w) + \frac{1}{1+r}\mathbb{E}_t\Big[&\big(\delta + (1-\delta)\mathbb{1}(S(x,y,z_{t+1})<0)\big)U(x,z_{t+1}) \\ &(1-\delta)\mathbb{1}(S(x,y,z_{t+1})\geq 0)\Big(s\lambda(\Gamma_{t+1})\int_0^1 \frac{v_t(y',\Gamma_{t+1})}{V(\Gamma_{t+1})}I(x,y,y',z_{t+1})dy' \\ &+(1-s\lambda(\Gamma_{t+1}))R(x,y,w,\Gamma_{t+1})\Big)\Big] \end{aligned} \tag{16}$$

with

$$I(x,y,y',z_t) = \begin{cases} S(x,y,z_t) & \text{if } S(x,y',z_t) > S(x,y,z_t) \geq 0 \\ S(x,y',z_t) & \text{if } S(x,y,z_t) > S(x,y',z_t) > 0 \\ 0 & \text{else} \end{cases} \tag{17}$$

and

$$R(x,y,w,\Gamma_t) = \begin{cases} W(x,y,\phi^0(x,y,\Gamma_t),\Gamma_t) & \text{if } \Delta(x,y,w,\Gamma_t) < 0 \leq S(x,y,z_t) \\ W(x,y,\phi^1(x,y,\Gamma_t),\Gamma_t) & \text{if } \Delta(x,y,w,\Gamma_t) > S(x,y,z_t) \geq 0 \\ W(x,y,w,\Gamma_t) & \text{else (status quo)} \end{cases} \tag{18}$$

The continuation value in equation (16) contains three components. The first line takes into account the probability that an unemployed worker looses her job, which could happen for exogenous or endogenous reasons. In case of a job loss next period, the worker receives the value of unemployment $U(x,z_{t+1})$. If the worker stays employed, two cases can occur. Either the worker meets with another firm, which is taken into account in the second line of equation (16). Or the no meeting occurs, but re-bargaining is still possible if one of the two parties has a credible threat to break the match. Intra-firm re-bargaining is taken into consideration in the third line of equation (16).

The function $I(.)$ captures wage changes following a meeting with an alternative employer. If the poaching firm, characterized by a productivity parameter y', is a better match for a worker of type x, the worker changes job and gets the full match surplus from its previous employer. If the poaching firm is not a better match for a worker of type x, the worker reveals the meeting to its current employer. The current employer makes a counter-offer that matches the best offer that firm y' can make. If the poaching firm is a not a credible employer, the meeting is not revealed and has no impact on the current match.

The function $R(.)$ takes into account intra-firm re-bargaining. If a worker has a credible threat to leave, the wage is re-bargained up to $\phi^0(x, y, \Gamma_t)$. This happens when the worker's surplus $\Delta(x, y, w, \Gamma_t) \equiv W(x, y, w, \Gamma_t) - U(x, z_t)$ is negative. When a firm's surplus $\Pi(x, y, w, \Gamma_t)$ is negative, the wage is re-bargained down to $\phi^1(x, y, \Gamma_t)$. The firm has a credible threat to break the match when the worker's surplus is greater than the joint surplus of a match $S(x, y, z_t) \equiv \Delta(x, y, w, \Gamma_t) + \Pi(x, y, w, \Gamma_t)$.

4.2 Dimension reduction

An important feature of equation (16) is that the endogeneous distribution of matches $h_t(x, y)$ only matters through next period's job meeting rate $\lambda(\Gamma_{t+1})$ and next period's endogenous distribution of vacancies $q(y, \Gamma_{t+1}) \equiv v(y, \Gamma_{t+1})/V(\Gamma_{t+1})$. To make a decision today, firms and workers only need to forecast 3 objects: next period's aggregate productivity level z_{t+1} (which is trivial given the $AR(1)$ assumption (6)) and the duo $(\lambda(\Gamma_{t+1}), q(y, \Gamma_{t+1}))$. To reduce the dimension of the relevant state variable to a finite dimensional object, let us follow a strategy similar to the one employed by Krusell and Smith (1998). Let us postulate that agents use a simple parametric forecasting rule to predict next period's job meeting rate:

$$\lambda_{t+1} = f_\lambda(\Omega|\boldsymbol{\theta}_\lambda) \tag{19}$$

with $\Omega \equiv (z_{t+1}, \lambda_t) \cup \Phi$, where Φ contains relevant variables known in the current period.[8] Agents also use a simple parametric rule to keep track of the endogenous distribution of vacancies $q(y, \Gamma_{t+1})$. As in Algan et al. (2008) or Winberry (2018), my strategy is to use a parametric function $q(y|\boldsymbol{q_{t+1}})$ to approximate $q(y, \Gamma_{t+1})$, where $\boldsymbol{q_{t+1}}$ is a finite-dimensional vector. In practice, I use a Beta density, which performs extremely well.[9] Agents are endowed with an additional forecasting rules to keep track of the shape parameters of the Beta density $\boldsymbol{q_{t+1}} = (a_{t+1}, b_{t+1})$:

$$\boldsymbol{q_{t+1}} = f_q(\Omega|\boldsymbol{\theta_q}) \tag{20}$$

Conditional on the forecasting rules $\boldsymbol{\theta} \equiv (\boldsymbol{\theta_\lambda}, \boldsymbol{\theta_q})$, the value of a y job to a worker of type x with wage w can be written as:

$$\begin{aligned} W(x, y, w, z_t, \lambda_t|\boldsymbol{\theta}) = w + \frac{1}{1+r}\mathbb{E}_t \Big[&\big(\delta + (1-\delta)\mathbb{1}(S(x, y, z_{t+1}) < 0)\big) U(x, z_{t+1}) \\ &(1-\delta)\mathbb{1}(S(x, y, z_{t+1}) \geq 0)\Big(s\lambda_{t+1} \int_0^1 q(y|\boldsymbol{q_{t+1}}) I(x, y, y', z_{t+1}) dy' \\ &+ (1 - s\lambda_{t+1}) R(x, y, w, \lambda_{t+1}, q(y|\boldsymbol{q_{t+1}}))\Big)\Big] \end{aligned} \tag{21}$$

The parameter values for the forecasting rules $\boldsymbol{\theta}$ can be estimated by Monte-Carlo by simulating an economy for a long period of time. Importantly, unlike in the model of Krusell and Smith (1998), the Monte-Carlo step and the calculation of the value function step are *independent* from each others. The time series needed to estimate $\boldsymbol{\theta}$, can be simulated without any reference to the value of a job $W(.)$. In a setting with heterogeneity and aggregate uncertainty, one generally has to find a fixed point for the value functions *and* the forecasting rules. In the present framework, because the forecasting rules and the value functions are orthogonal to each others, $W(.)$ has to be calculated only *once*. Besides the computational edge of the present setting, I see the main advantage of the current setting as a theoretical one. Conditional on the forecasting rules $\boldsymbol{\theta}$, the Bellman operator implicitly defined in equation (21) is a contraction. Hence, $W(.)$ exists and is

[8] For instance, Φ may contain z_t, the square of λ_t and an interaction term between λ_t and z_t.

[9] Using a histogram to approximate $q(y, \Gamma_{t+1})$, as in Reiter (2009), is an attractive alternative. Agents would forecast the value of $q(y, \Gamma_{t+1})$ on a deterministic grid with N elements: $\boldsymbol{q_{t+1}} = (q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y)) = f_q(\Omega_t|\Theta_q))$. If the economy features N discrete firm types instead of a continuum, this approach amounts to assuming that agents forecast the number of vacancies posted by each firm type. I discuss this alternative in the Appendix.

unique.

In practice, I build Ω so that it contains $(z_{t+1}, z_t, \lambda_t)$, their square and interactions terms. I use the LASSO to determine which variables are important to forecast next period's job meeting rate. A simple forecasting rule with only first order terms emerges from this procedure:

$$\lambda_t = -0.141 + 0.675\lambda_{t-1} + 0.194 z_t$$

This result is reminiscent of the *approximate aggregation* finding, prevalent in macroeconomic models with heterogeneous agents. While a priori the state variable is infinite dimensional, once individual policy rules are aggregated, a simple rule emerges from complexity. Using the LASSO generates forecasting rules that are robust to over-fitting, as measured by the out-of-the-sample maximum absolute percentage error. A series of accuracy tests is reported in the Appendix (see section C.1).

4.3 Evolution of wages

Let us introduce the notation $\hat{\Gamma}_t \equiv (z_t, \lambda_t | \hat{\boldsymbol{\theta}})$, denoting the approximate aggregate state variable (conditional on the forecasting rule $\hat{\boldsymbol{\theta}}$). At every period t, each firm y offers three types of wages for each worker of type x. A starting wage $\phi^0(x, y, \hat{\Gamma}_t)$ is offered to worker moving out of unemployment, or when the worker's surplus gets negative (and the match is still feasible). The starting wage $\phi^0(x, y, \hat{\Gamma}_t)$ is implicitly defined by

$$W(x, y, \phi^0(x, y, \hat{\Gamma}_t), \hat{\Gamma}_t) = U(x, z_t) \tag{22}$$

When the firm's surplus is negative (and the match is still feasible), the wage is rebargained down to $\phi^1(x, y, \hat{\Gamma}_t)$. This wage is implicitly defined by

$$W(x, y, \phi^1(x, y, \hat{\Gamma}_t), \hat{\Gamma}_t) = S(x, y, z_t) \tag{23}$$

When an employed worker contacts another firm, the resulting wage, denoted by $\phi^2(x, y, y', \hat{\Gamma}_t)$, depends on the incumbent and the poaching firms:

$$\phi^2(x, y, y', \hat{\Gamma}_t) = \begin{cases} \phi^1(x, y', \hat{\Gamma}_t) & \text{if } S(x, y, z_t) > S(x, y', z_t) > 0 \\ \phi^1(x, y, \hat{\Gamma}_t) & \text{if } S(x, y', z_t) > S(x, y, z_t) \geq 0 \end{cases} \tag{24}$$

If a worker x, currently working with firm y, meets with y' (a less productive match compared to firm y), firm y responds by offering the maximum wage firm y' can offer: $\phi^1(x, y', \hat{\Gamma}_t)$. If firm y' is a better match for worker x, firm y' offers a wage such that firm y's maximum offer is not enough to retain worker x.

4.4 Distribution of (starting) wages

The flow equation for the joint evolution of matches and wages, denoted by $h_t(x, y, w)$, is complicated since it involves its past values (see the Appendix). It is also cumbersome to approximate $h_t(x, y, w)$ since the wage dimension is inherently continuous. One could use the method developed in Young (2010) to approximate $h_t(x, y, w)$, or use a panel with a sufficiently high number of agents. When estimating the model, I use a much simpler endogenous object: the distribution of *starting wages* for workers exiting unemployment, denoted by $h_{t,0}(x, y) \equiv h_t(x, y, \phi^0(x, y, \hat{\Gamma}_t))$. Each period, the equation for $h_{t,0}(x, y)$ solves:

$$h_{t,0}(x, y) = u_{t+}(x)\lambda_t \frac{v(y, \hat{\Gamma}_t)}{V(\hat{\Gamma}_t)} \mathbb{1}\{S(x, y, z_t) \geq 0\} \tag{25}$$

Contrary to the joint distribution of matches and wages, $h_{t,0}(x, y)$ is only bi-dimensional and is memory-less, which makes it an appealing object for estimation purposes.

The model predicts interesting wage dynamics. On the one hand, when the aggregate productivity parameter is high, workers expect it to stay high in the future. Higher future prospects decrease today's reservation wage, putting a downward pressure on the starting wage ("expectation effect"). The expectation effect is particularly strong in the present setting because workers have zero bargaining power. On the other hand, today's output value $p(x, y, z)$ goes up, increasing the value of workers from the firms' perspective, making them more willing to pay high wages ("output value effect").

These two effects are combined with the dynamics of sorting along the cycle, which may alter the dynamics of aggregate wages through composition effects. A "cleansing" effect implies that only the better matched workers stay employed during a downturn. Simultaneously, firms may find it harder to employ good workers in recessions, driving them to post more low-quality jobs (the "sullying effect", see Barlevy (2002)).

5 Parametrization and Estimation

A key ingredient to jointly match wage and employment moments is to allow for home production $b(.)$ to be a function of the aggregate state variable z_t. A tendency of $b(x, z_t)$ to be increasing with z would limit the "expectation effect", which pushes down starting wages in good times. Why would home production depend on z in the first place? One leading reason is that home production contains unemployment benefits, which are generally equal to a fraction of past labor income. Empirically, labor income tends to increase in booms, causing unemployment benefits to rise. I choose a parametrization that nests the one in Lise and Robin (2017).[10] More specifically, I use a home production of the form:

$$b(x,z) = (b_0 + b_1 z) \times \bar{b}(x) \tag{26}$$

where $\bar{b}(x) \equiv 0.7 \times p(x, y^*(x), 1)$ and $y^*(x) = \arg\max_y p(x, y, 1)$. The function $\bar{b}(x)$ captures a fixed proportion of output when the worker is in his/her optimal match, at the neutral technological state. The aggregate productivity level follows an $AR(1)$ process of the form:

$$z_{t+1} = \rho z_t + \sigma\sqrt{1-\rho^2}\varepsilon_{t+1} \tag{27}$$

Worker's ability types x are distributed according to a Beta distribution with shape parameters β_1 and β_2. Output at the match level is given by a polynomial of the form:

$$p(x,y,z) = z(p_1 + p_2 x + p_3 y + p_4 x^2 + p_5 y^2 + p_6 xy) \tag{28}$$

I assume the following simple form for the cost of posting v vacancies:

$$c(v) = c_0 \frac{v^{1+c_1}}{1+c_1} \tag{29}$$

I assume a Cobb-Douglas function for the matching function:

$$M_t = \alpha L_t^{\omega} V_t^{1-\omega} \tag{30}$$

[10] I do this for both theoretical and practical reasons. From a theoretical point of view, it is reassuring that the parametrization can simplify to one that was proven to successfully replicate a vast array of employment moments. From a practical perspective, it is convenient that my parametrization nests an existing one, as it helps me in formulating a meaningful set of starting values.

A likelihood-based estimation is unfeasible because the likelihood function is not tractable. Instead, I estimate the model using the Simulated Method of Moments (SMM) with US data. I include employment-related moments and moments from the starting wage distribution. In practice, I include the elasticity of the first 9th deciles of the starting wage distribution in the list of moments to be matched (see Tables 8 and 9). In total, I use 35 moments to estimate 16 parameters. Practical details regarding the estimation procedure are listed in the section G of the Appendix .

5.1 Estimated values and model fit

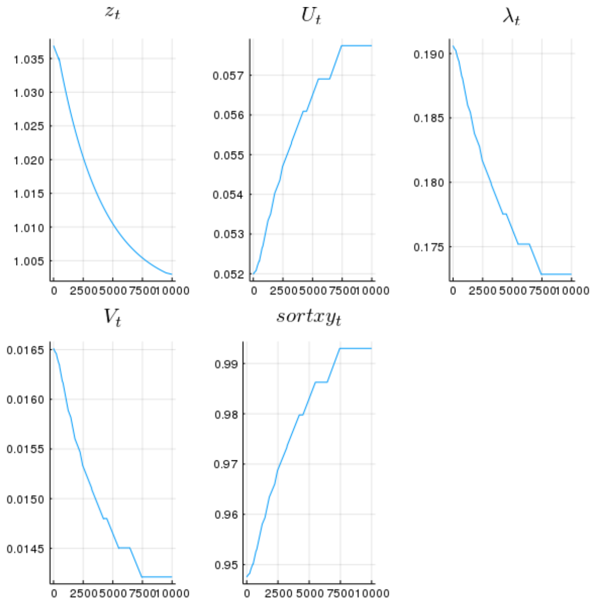
Estimated parameters are presented in Table 7. The parameter b_1 is found to be positive. Indexation of unemployment benefits on past labor income would imply this feature. Because home production is slightly pro-cyclical, unemployed workers are more picky during expansions. Having more selective workers generates mitigates the "expectation effect" on the worker side, potentially pushing down the reservation wage. The impulse response function indicates that the model features a "cleansing effect", with sorting quality between firms and workers declining by approximately 5% after a one-standard deviation productivity shock (Figure 2). The parameter estimates imply that the labor market features associative matching, with high productivity firms preferring to match with high productivity firms (see Figure 3). The value for the parameter s suggests that unemployed workers are searching for a job with an intensity that is 16 times bigger than the intensity of already employed workers.

Table 7: Estimated Parameters

Parameter	Value	Estimated	Description
r	0.05	No	annual interest rate
α	0.501	Yes	matching function parameter
ω	0.5	No	matching function parameter
s	0.061	Yes	search intensity parameter
c_0	0.405	Yes	vacancy posting cost parameter
c_1	0.030	Yes	vacancy posting cost parameter
δ	0.012	Yes	exogenous separation rate
ϕ	0.083	Yes	productivity shock parameter
ρ	0.999	No	productivity shock parameter
β_1	13.490	Yes	worker heterogeneity parameter
β_2	16.735	Yes	worker heterogeneity parameter
p_1	0.053	Yes	value added parameter
p_2	2.162	Yes	value added parameter
p_3	-0.157	Yes	value added parameter
p_4	8.818	Yes	value added parameter
p_5	-1.880	Yes	value added parameter
p_6	7.126	Yes	value added parameter
b_0	0.478	Yes	home production parameter
b_1	0.814	Yes	home production parameter

Notes: Parameter values were estimated using the Simulated Method of Moments. Parameter values were rounded to the nearest thousandths. I take the value of parameter ρ, characterizing the persistence of TFP (at the weekly frequency) from Lise and Robin (2017).

Figure 2: Impulse response function after a productivity shock

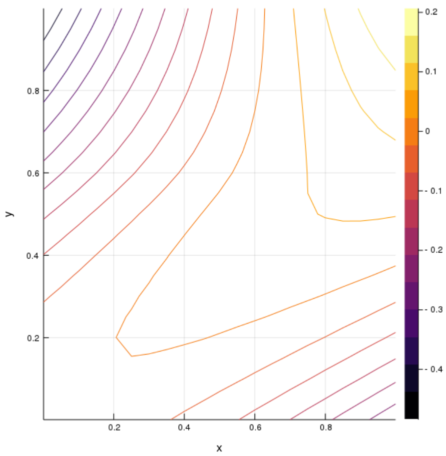


Notes: This figure shows the response of the unemployment rate U_t, the job meeting rate λ_t, the number of vacancies V_t, and sorting after a one-standard deviation positive TFP shock. I quantify sorting using the formula

$$sortxy_t = \frac{1}{C} \exp\Big(- \int_0^1 \int_0^1 h_t(x,y)\big(y - y^*(x,z_t)\big)dxdy\Big)$$

where $y^*(x,z)$ is the firm type that would maximize the joint surplus of a match for a worker of type x when the productivity parameter is equal to z_t. C is a normalizing constant chosen such that the best observed sorting value is equal to 1. A value of $sortxy_t$ below 1 indicates a sub-optimal pairing between firms and workers, in the sense that production could be improved by reallocating workers to firms with productivity types closer to $y^*(x,z_t)$.

Figure 3: Value of net production at the match level $s(x,y,1)$



Notes: This figure shows a contour plot for the value of net production at the match level $s(x,y,1) = p(x,y,1) - b(x,1)$ when the aggregate state variable z_t is at its neutral state ($z_t = 1$).

Table 8: Elasticity of wage and starting wage deciles

	New hires		All workers	
	Data	Model	Data	Model
P10	0.44	0.45	0.20	0.40
P20	0.32	0.38	0.13	0.32
P30	0.34	0.35	0.18	0.33
P40	0.31	0.34	0.18	0.30
P50	0.35	0.35	0.12	0.28
P60	0.42	0.40	0.30	0.31
P70	0.37	0.42	0.26	0.32
P80	0.45	0.68	0.28	0.39
P90	0.65	0.77	0.46	0.55

Notes: This table displays the elasticity of starting wage deciles with respect to changes in aggregate labor productivity. The first column is based on the CEPR CPS ORG dataset and own calculations. The second column is based on simulated data generated using the my novel parametrization. The fourth column is based on simulated data generated using a sample of 1000 agents during 6000 periods (weeks), discarding the first 1000 periods.

Table 9: Empirical and Simulated Employment Moments. US (1951-2012)

	Data	Model		Data	Model
$E[U]$	0.058	0.059	$std[U_{27+}]$	0.478	0.150
$E[U_{5+}]$	0.035	0.028	$std[UE]$	0.127	0.089
$E[U_{15+}]$	0.018	0.010	$std[EU]$	0.100	0.058
$E[U_{27+}]$	0.010	0.006	$std[EE]$	0.095	0.116
$E[UE]$	0.421	0.527	$std[V/U]$	0.381	0.333
$E[EU]$	0.025	0.032	$corr[U,V]$	-0.846	-0.985
$E[EE]$	0.025	0.015	$corr[U,VA]$	-0.860	-0.983
$E[V/U]$	0.634	0.251	$corr[V,VA]$	0.721	0.994
$std[V]$	0.206	0.188	$corr[UE,VA]$	0.878	0.957
$std[VA]$	0.033	0.046	$corr[EU,VA]$	-0.716	-0.970
$std[U]$	0.191	0.146	$corr[UE,EE]$	0.695	0.947
$std[U_{5+}]$	0.281	0.236	$autocorr[GDP]$	0.932	0.987
$std[U_{15+}]$	0.395	0.257			

Notes: Data columns are from Lise and Robin (2017). $E[U]$ is the average quarterly unemployment rate. $E[U_{5+}]$, $E[U_{15+}]$ and $E[U_{27+}]$ are the average quarterly unemployment rates for more than 5, 15 and 27 weeks respectively. $E[UE]$, $E[EU]$ and $E[EE]$ are the average quarterly job-finding, job-losing and job-to-job transition rates. V denotes the number of vacancies, and VA is value added. $std[x]$ denotes the standard deviation of the variable x. $corr[x,y]$ denotes the correlation between variables x and y. $autocorr[x]$ denotes the auto-correlation of variables x.

5.2 Identification

I provide a heuristic justification of my identification strategy and I rely on numerical tools to back up my intuitions. First of all, exit from unemployment and job-to-job mobility are key in determining the value of search efficiency α and the relative search intensity between unemployed and unemployed worker s. A higher job-finding rate should be associated with a higher value of and more job-to-job transitions should indicate a higher value for s. The steady-state value (or its long-run average) of the unemployment rate is informative on the exogenous job-destruction rate δ. The unemployment rate for different duration and its volatility are informative on the distribution of types within the economy. More long-term unemployment indicates a distribution of types tilted towards low types. The elasticity of starting wages captures to what extent the "cleansing" effect (survival of better workers) and the "sullying" effect (firms posting low quality jobs) dominate during recessions. Thus, deciles of the starting wage distribution provide valuable information on the matching function, the vacancy cost function, the distribution of types and parameters for $b(.)$ and $p(.)$. Table 15, reporting values for the Jacobian of the function mapping parameters to simulated moments $f : \boldsymbol{p} \rightarrow \boldsymbol{m}$, largely confirms these intuitions. In addition to the statistics discussed above, market tightness is found to be key in disciplining the model. This not surprising given that the market tightness contains information on both sides of the market. Interestingly, the partial derivatives of f with respect to the deciles of the wage distribution are one or two order of magnitude larger the partial derivatives involving employment moments. The information contained in wages is substantial. More complex parametrization could potentially be estimated using the extra information contained in the variation of wages.

6 Labor income shocks and sorting

What are the main mechanisms behind labor income shocks? In particular, what is the contribution of the inter-firm channel to variations over the business cycle? To quantify the importance of the inter-firm channel on labor income shocks, I first measure the persistence of labor income losses for displaced workers. If follow the literature on displaced workers (see for instance Stevens (1997)) and run the following regression on simulated data:

$$y_{it} = \sum_{k=0}^{10} \delta_k D_{it}^k + \varepsilon_{it} \tag{31}$$

The variable D_{it}^k is an indicator variable equal to 1 if worker i was a displaced worker k periods ago. The coefficients δ_k measures the current effect of job displacement on y_{it}. I classify a worker as displaced if the worker experienced unemployment in year k for at least 10 weeks and if the worker was poorly matched upon finding a new job. I consider a worker to be poorly matched if the (absolute value) of the distance between the current firm y and the optimal match y^* is bigger than a threshold value [11]. The left panel of Figure 4 displays the recovery of yearly earnings, the number of weeks worked and the hourly wage for displaced workers relative to non-displaced workers using simulated data. Displaced workers experience on average a 37% drop in yearly labor income the year of displacement, mainly driven by a loss in weeks worked. As the number of weeks worked quickly recovers, the milder initial loss in terms of hourly wage takes much longer to recover. Long-lasting labor earnings losses are driven by mismatches after a job loss. As sorting improves over time, the initial loss in hourly wage slowly vanishes. Because search for better firm types is random, improvements in sorting takes time to materialize. These results are consistent with the empirical patterns reported by the literature on job displacement. The right panel of Figure 4, based on empirical work from Huckfeldt et al. (2016), also shows that the initial impact of a job loss is mainly driven by a loss in annual hours worked. However, as predicted by the model, the persistence in labor earning losses is to be attributed to an enduring loss in wages. While the model predicts that displaced workers eventually recover from a job loss, the right panel of Fig-

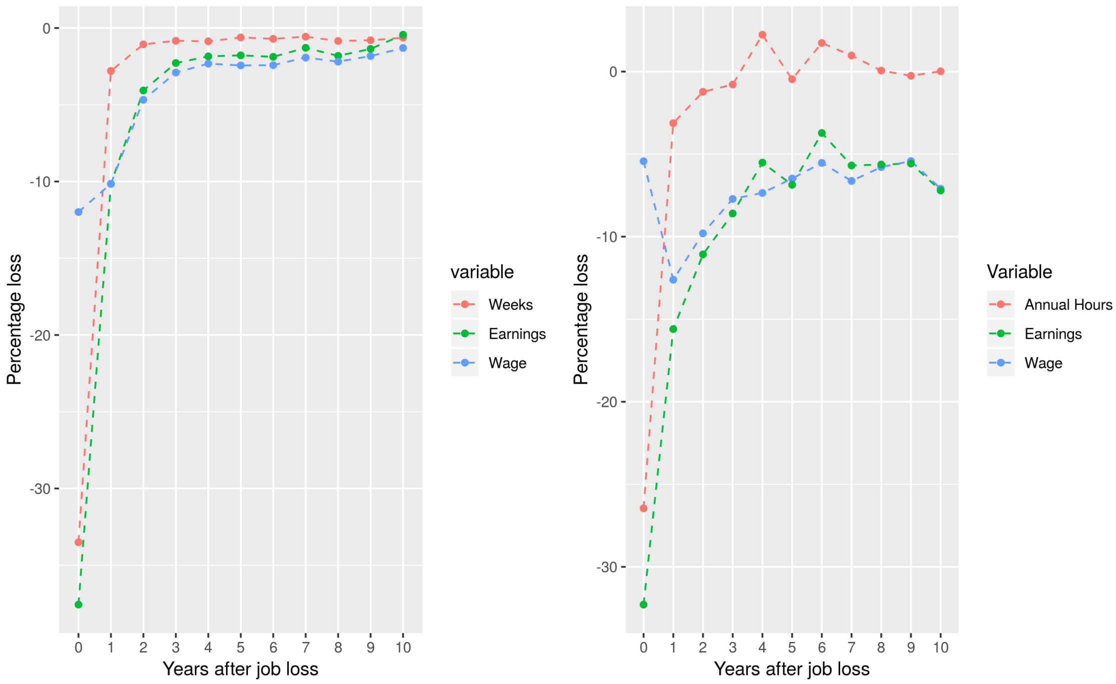
[11] As a threshold value, I use 2 times the standard deviation of the distances (absolute value difference) between the current firm types and the optimal firm type y^*: $|y - y^*|$)

ure 4 indicates that empirically there is a 5% permanent drop in wages. The permanent drop in wages can be explained by a loss in human capital accumulation for displaced workers, which is not included in the present model.

To further understand how labor income shocks and sorting are related, I run an experiment in which I keep only one type of firm in the economy (the median firm). Results are presented in Table 10. When the inter-firm channel is nonexistent, business cycle fluctuations in yearly labor income are reduced by approximately 12% compared to the baseline model with firm heterogeneity. The inter-firm channel is particularly important for the tails of labor income changes. When the inter-firm channel is shut down, the *change* over the business cycle in the probability of experiencing more than a 50% drop in labor income over a year is reduced by approximately 19%. The change over the business cycle in the likelihood of experiencing more than a 50% increase in labor income over a year is reduced by 97.5%.

Why does sorting matter for changes in labor income over the cycle? In the present model, changes in labor income are either caused by a change in employment status (employed or unemployed) or by a change in wages. Changes in wages are driven by the inter-firm channel. That is, the reallocation of workers from bad matches to better matches. The inter-firm channel depends on the number of meetings occurring each period. In a recession, the number of on-the-job meeting $s\lambda_t$ plummets causing the inter-firm channel to dry up. As a result, in the after-math of a bad productivity shock, workers stay mismatched for longer period of time generating labor income losses. In an economy with only one type of firm, this dimension vanishes and the model generates more modest fluctuations in labor along the business.

Figure 4: Displaced workers and income losses



(a) Model

(b) Data

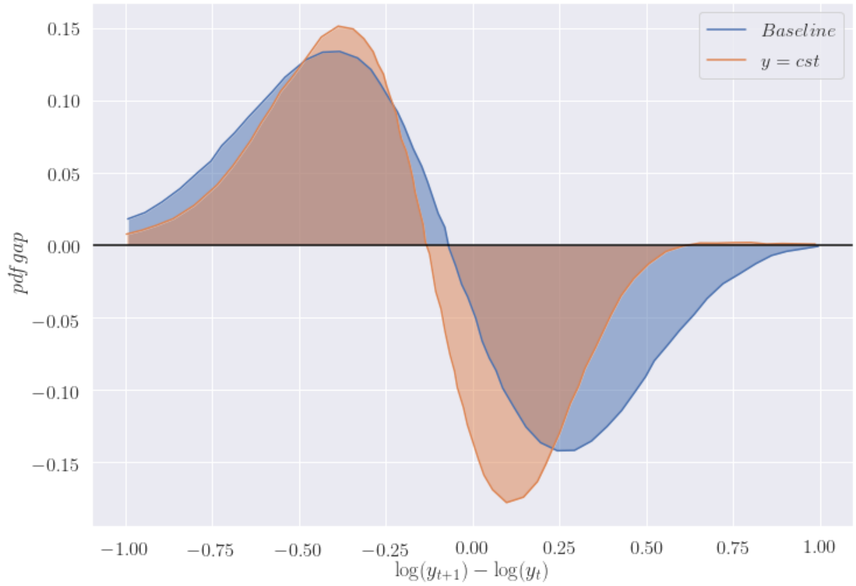
Notes: The left panel shows the effect of job displacement on yearly labor income, the annual number of weeks worked, and the hourly wage relative to workers who have not been displaced within the past ten years. Calculations are based on simulated data. The right panel is from Huckfeldt et al. (2016), based on data from the Panel Study of Income Dynamics from 1968 to 1997. It shows the (relative) effect of job displacement on annual hours worked, yearly labor income, and hourly wage.

Table 10: Variations in the probability of labor income changes over the cycle

	Changes in probabilities over the cycle		
	Data	Baseline	$y = cst$
A	0.164	0.149	0.133
$\Delta \Pr(\Delta Income_{it} < 50\%)$	1.3%	3.1%	2.5%
$\Delta \Pr(\Delta Income_{it} > 50\%)$	-1.5%	-1.4%	-0.035%
$\Delta \Pr(\Delta Income_{it} < 25\%)$	2.9%	6.3%	5.5%
$\Delta \Pr(\Delta Income_{it} > 25\%)$	-3.0%	-4.4%	-1.8%

Notes: The first column is based on data from Guvenen et al. (2014) (US data, 1995–96 versus 2008–9.). The second and third columns are based on simulated data. The variable A is defined as the area between the line representing the density gap and the x-absis. the pdf gap is defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null and A would be equal to 0 (see Figure 9). Other rows measure the *change* in the probability of a certain event relative to yearly labor income. For instance, the row $\Delta \Pr(\Delta Income < 50\%)$ measures the *change* in the probability that a worker loses more than 50% of her labor income over a year (recession minus expansion). The second column is based on simulated data from the model with the estimated parameters. The third column is based on simulated data with only one type of firm (the median firm).

Figure 5: Change in labor income risk over the cycle and sorting



Notes: This figure shows the density gap, defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null. The blue line is the baseline model. The orange line is based on a simulation with only one type of firm (the median firm).

7 Optimal UI

The analysis has established that sorting between firms and workers is central for idiosyncratic labor income risk, especially when considering extreme labor income changes. Can a government improve the market outcome using a simple unemployment benefit rule? This question has already been explored by Lise et al. (2016) in a similar setting, but without aggregate uncertainty. The authors find that the optimal unemployment scheme delivers an improvement of 1.4%. One may wonder to what extent their findings extend to a model with aggregate risk. One may also wonder to what extent the government may attenuate fluctuations in labor income with a simple unemployment benefit scheme. Before analyzing the optimal unemployment benefit policy, what are the imperfections that would justify government intervention in the first place? In a model featuring matching and search frictions, there exist congestion externalities (Hosios (1990)). Some types of workers may be searching too much, especially in the present setting in which workers and firms are heterogeneous. Some high-type firms would probably post more vacancies if they were certain to find only high-type workers on the labor market. For this reason, it might be optimal that low-type workers search less, which could be incentives by providing a higher replacement rate. I solve for an optimal unemployment benefit equal to a fix proportion of the expected labor income at the steady-state:

$$b_{UI}(x) = b \int_0^1 p(x, y, 1) h(y|x) dy \tag{32}$$

The unemployment insurance is funded by a proportional tax on match output $\int b_{UI}(x)u(x)dx = \tau \int p(x, y, 1)h(x, y)dxdy$. The welfare criterion I use is the sum of market output (net of taxes), plus home production and UI, minus the cost of creating vacancies:

$$\begin{aligned} W_{b_0} = & \int_0^1 \int_0^1 (1 - \tau) h(x, y) p(x, y, 1) dxdy + \int_0^1 \big(b(x, 1) + b_{UI}(x)\big) u(x) dx + \\ & - \int_0^1 c(v(y)) dy \end{aligned} \tag{33}$$

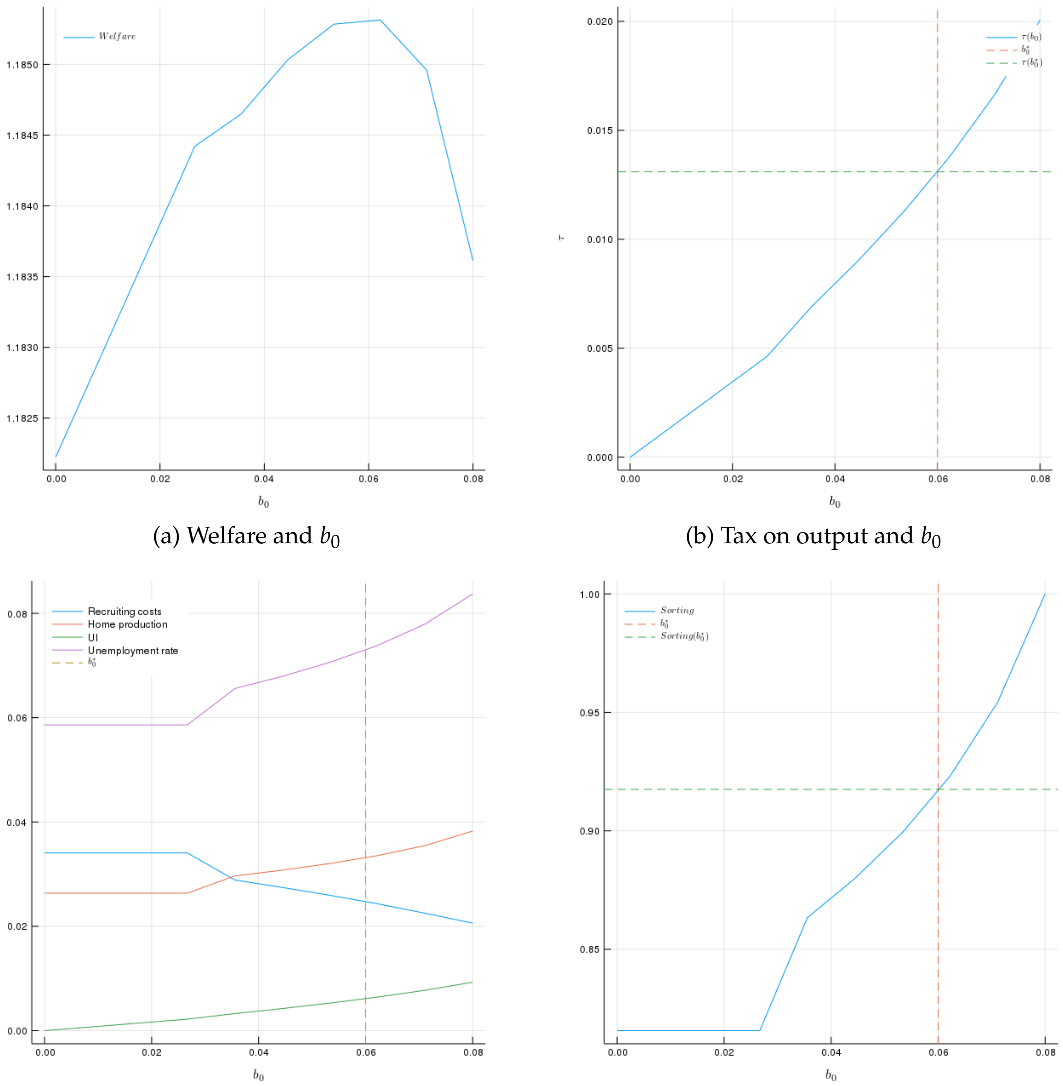
Equation (33) underlines the mechanism at play. Taxing output hurts employed workers and improves the welfare of unemployed workers. Yet, the composition of matches is altered and firms may change the amount of vacancies they post. Results are presented in Figure 6. The optimal value for b_0 is approximately 0.06, which is funded by a 1.3% tax on output at the match level. This is slightly higher than the 0.95% tax rate found in Lise et al. (2016). This tax on output is used to fund an unemployment insurance that represents approximately 19.2% of the aggregate value of home production. The unemployment insurance scheme rises the unemployment rate to approximately 7%. The welfare gains are driven by a 13.16% increase in sorting and a decrease in the cost of creating vacancies, which are enough to offset the distortions created by the tax on output. Table 11 underlines that UI stabilizes labor income by approximately 2%. This policy fosters employment for high type workers and firms. The tax on output provides a safety net for low-skilled workers more likely to bear the burden of unemployment along the cycle.

Table 11: Variations in the probability of labor income changes over the cycle with the optimal UI

	Changes in probabilities over the cycle		
	Data	Baseline	UI
A	0.164	0.149	0.146
$\Delta P(\Delta Income_{it} < 50\%)$	1.3%	3.1%	3.1%
$\Delta P(\Delta Income_{it} > 50\%)$	-1.5%	-1.4%	-1.2%
$\Delta P(Income_{it} < 25\%)$	2.9%	6.3%	6.2%
$\Delta P(Income_{it} > 25\%$	-3.0%	-4.4%	-4.2%

Notes: The first column is based on data from Guvenen et al. (2014) (US data, 1995–96 versus 2008–9.). The second and third columns are based on simulated data. The variable A is defined as the area between the line representing the density gap and the x-absis. the pdf gap is defined as the pdf for $(\log(y_{t+1}) - \log(y_t))$ in recession minus the pdf for $(\log(y_{t+1}) - \log(y_t))$ in expansion. If the business cycle did not generate changes in labor income, the density gap would be null and A would be equal to 0. Other rows measure the *change* in the probability of a certain event relative to yearly labor income. For instance, the row $P(Income < 50\% decrease)$ measures the *increase* in the probability that a worker loses more than 50% of her labor income over a year when the economy is in recession. The second column is based on simulated data from the model with the estimated parameters. The third column is based on simulated data with only one type of firm (the median firm).

Figure 6: Optimal UI



(a) Welfare and b_0

(b) Tax on output and b_0

(c) Production and b_0

(d) Sorting and b_0

Notes: The top-left quadrant shows the link between the replacement rate b_0 and welfare. Welfare is the sum of net market production, home production, the total value of UI, minus the cost of creating vacancies. The top right-quadrant shows the link between b_0 and the corresponding tax rate on production τ. The bottom-left quadrant shows the link between b_0 and selected elements of the welfare function. The bottom-right quadrant shows the link between b_0 and an index measuring the quality of sorting between firms and workers.

8 Conclusion

This paper analyzes the determinants of labor income changes over the business cycle. The novelty in this analysis is to show that sorting between firms and workers matters when considering fluctuations in labor income. The mechanism is quite intuitive. Because of search frictions on the labor market, workers and firms are not necessarily well matched. The pairing between firms and workers is improved by the slow process of job-to-job transitions. In a recession, less vacancies are posted and the inter-firm channel dries up. Workers accumulate labor income losses by working with firms that are not optimal for them. While the primary driver of income losses in recession is unemployment, the sorting channel accounts for 12% of fluctuations in labor income. A simple policy can generate welfare gains in that context: unemployment insurance. By varying the replacement rate, the government alters incentives for different types of workers. In particular, low-skilled workers are more patient. High-skilled workers benefit from the resulting reduction in congestion effects. Improved sorting is more than enough to offset the distortion effects created by taxing output.

To arrive to this conclusion, I developed and estimated a dynamic search-and-matching model with heterogeneous firms and workers. While a priori not tractable, the model simplifies considerably by realizing that the state variable can be reduced to a finite dimensional vector, without losing much generality nor stability. The key to tractability lies in the fact that employment flows can be determined in a first step, independently from the wage allocation problem. While wages do dependent on next period's job meeting rate and distribution of jobs across firm and worker types, as it is generally the case in dynamic search-and-matching models with heterogeneity, dimension reduction techniques can be used. Because the simulated series needed to perform the dimension reduction step are independent from the wages, the value function characterizing to the wage problem exists and is obtained immediately, without the need of an outer loop.

The clear separation of the employment problem from the determination of wages results from the combination of three assumptions: zero bargaining power for unemployed workers, determination of wages according to the sequential auction framework and linear utility. The existence of an other set of assumptions implying a similar separation between employment flows and wages is still an open question. In particular, relaxing the assumption of linear utility is important to analyze precautionary savings for this class of models. The extent to which the conclusions of this paper can be extended to a framework with risk-averse workers, having access to a risk-free asset, is currently being investigated.

References

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

Yann Algan, Olivier Allais, and Wouter J Den Haan. Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, 32(3):875–908, 2008.

Yann Algan, Olivier Allais, Wouter J Den Haan, and Pontus Rendahl. Solving and simulating models with heterogeneous agents and aggregate uncertainty. In *Handbook of Computational Economics*, volume 3, pages 277–324. Elsevier, 2014.

Adrien Auclert, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. Technical report, National Bureau of Economic Research, 2019.

Gadi Barlevy. The sullying effect of recessions. *The Review of Economic Studies*, 69(1):65–96, 2002.

Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739, 2019.

Timo Boppart, Per Krusell, and Kurt Mitman. Exploiting mit shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89:68–92, 2018.

Darrell Duffie and Kenneth J Singleton. Simulated moments estimation of markov models of asset prices, 1990.

Christian Gourieroux, Alain Monfort, and Eric Renault. Indirect inference. *Journal of applied econometrics*, 8(S1):S85–S118, 1993.

Fatih Guvenen, Serdar Ozkan, and Jae Song. The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660, 2014.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of us workers reveal about life-cycle earnings risk? Technical report, National Bureau of Economic Research, 2015.

Christian Haefke, Marcus Sonntag, and Thijs Van Rens. Wage rigidity and job creation. *Journal of monetary economics*, 60(8):887–899, 2013.

Arthur J Hosios. On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies*, 57(2):279–298, 1990.

Christopher Huckfeldt et al. Understanding the scarring effect of recessions. *Report, Economics Department*, 2016.

Per Krusell and Anthony A Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896, 1998.

François Le Grand, Xavier Ragot, et al. Optimal fiscal policy with heterogeneous agents and aggregate shocks. *Document de travail*, 2017.

Jeremy Lise and Jean-Marc Robin. The macrodynamics of sorting between workers and firms. *American Economic Review*, 107(4):1104–35, 2017.

Jeremy Lise, Costas Meghir, and Jean-Marc Robin. Matching, sorting and wages. *Review of Economic Dynamics*, 19:63–87, 2016.

Jeremy Lise, Jean Marc Robin, Rasmus Lentz, et al. The macrodynamics of the wage distribution. In *2017 Meeting Papers*, number 1220. Society for Economic Dynamics, 2017.

Bartosz Mackowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009.

Guido Menzio and Shouyong Shi. Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4):1453–1494, 2010.

Min Ouyang. The scarring effect of recessions. *Journal of Monetary Economics*, 56(2):184–199, 2009.

Jonathan A Parker and Annette Vissing-Jorgensen. Who bears aggregate fluctuations and how? *American Economic Review*, 99(2):399–405, 2009.

Christopher A Pissarides. The unemployment volatility puzzle: Is wage stickiness the answer? *Econometrica*, 77(5):1339–1369, 2009.

Fabien Postel-Vinay and Helene Turon. On-the-job search, productivity shocks, and the individual earnings process. *International Economic Review*, 51(3):599–629, 2010.

Xavier Ragot and François Le Grand. Managing inequality over the business cycles: Optimal policies with heterogeneous agents and aggregate shocks. Technical report, 2019.

Michael Reiter. Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665, 2009.

Jean-Marc Robin. On the dynamics of unemployment and wage distributions. *Econometrica*, 79(5):1327–1355, 2011.

Edouard Schaal. Uncertainty and unemployment. *Econometrica*, 85(6):1675–1721, 2017.

Robert Shimer. The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 95(1):25–49, 2005.

Jae Song, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. Firming up inequality. *The Quarterly journal of economics*, 134(1):1–50, 2018.

Ann Huff Stevens. Persistent effects of job displacement: The importance of multiple job losses. *Journal of Labor Economics*, 15(1, Part 1):165–188, 1997.

Karl Walentin and Andreas Westermark. Learning on the job and the cost of business cycles. 2018.

Thomas Winberry. A method for solving and estimating heterogeneous agent macro models. *Quantitative Economics*, 9(3):1123–1151, 2018.

Eric R Young. Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36–41, 2010.

A Numerical details

When estimating the model, I choose a rather coarse grid for the x and y dimensions using 11 points. Given that these dimensions represent non-time varying discrete types, this is done without loss of generality. When solving for the functions $U(.)$, $S(.)$ and $W(.)$, I treat z as a continuous dimension and I use Gaussian quadrature nodes. To produce results comparable with Lise and Robin (2017), I use an evenly-spaced discrete grid for z when simulating time series, with 51 different values for z. Regarding the forecasting rule $f(\Omega|\boldsymbol{\theta})$, Ω contains a constant, $log(z_t)$, $log(z_{t-1})$, λ_t, λ_{t-1}, their square and interactions terms. I use the LASSO to estimate the forecasting rule, which has the advantage of automatically selecting the relevant variables needed to make accurate and reliable predictions. [12] I solve $W(x, y, z, \lambda|\boldsymbol{\theta})$ by value function iteration. To calculate moments on the full distribution of wages, I use a panel of 1000 agents. I simulate the panel for 6000 periods (weeks) and I drop the first 1000 observations to get rid of the potential impact of initial values. To minimize the SMM objective function, I use a parallel multi-start approach. I start several Nelder-Mead algorithms in parallel with different starting values. The global minimum is the minimum of the several minima for which convergence was reached. The code is implemented in Julia 0.6.4 (see Bezanson et al. (2017)).

B Algorithm to solve and simulate the model

To solve the model, I proceed as follows:

1. Solve S(.) and U(.) by value function iteration, as in Lise and Robin (2017).

2. Simulate an economy for a long period of time and discard the first 10th observations. This step can be achieved independently of $W(.)$. This generates a synthetic sample containing $(z_t, \lambda_t)_{t=1}^{T}$ and any other variables of interest.

[12] For the LASSO, I use the package GLMNet.jl

3. Find an estimate of the parametric forecasting rule, denoted by $\hat{\boldsymbol{\theta}}$.

4. Find a solution to $W(x, y, w, \hat{\Gamma}_t|\hat{\boldsymbol{\theta}})$ by value function iteration.

5. Using $W(x, y, w, \hat{\Gamma}_t|\hat{\boldsymbol{\theta}})$, $U(x, z_t)$ and $S(x, y, z_t)$, calculate the value of ϕ^0, ϕ^1 and ϕ^2 for every t of the synthetic data.

C Accuracy of approximations

My computational strategy relies on two approximations: (i) next period's job meeting rate and parameters of $\tilde{q}(y|\boldsymbol{q_t})$ are well predicted by simple forecasting rules (ii) the endogenous distribution of vacancies $q(y)$ can reasonably be approximated by a parametric function. This section shows that both approximations are accurate. This section also shows that one may dispense with the parametric assumption for the distribution of vacancies and instead use a histogram.

C.1 Forecasting rules accuracy

I calculate how well forecasting rules predict the paths for λ_t and $\boldsymbol{q_t}$ within the sample that was used to estimate the forecasting rules (within-sample prediction power). I also simulate a new sample and I compare the actual realization of time series to the predictions implied by the forecasting rule (out-of-the-sample prediction power). The within-sample accuracy can be visualized in Figure . The estimated forecasting rules are quite precise, as underlined in Table 14. For instance, the median percentage error for the job meeting rate is only 0.13%. Is this number within a credible range? The literature on rational inattention has described reasons why firms may optimally commit small errors. The full optimization problem may too hard or too resource-consuming to solve. Firms may decide to use only a fraction of the full information set each period (Mackowiak and Wiederholt (2009))). The estimated forecasting rule for the job meeting rate is the following simple linear relationship:

$$\lambda_t = -0.141 + 0.675\lambda_{t-1} + 0.194z_t$$

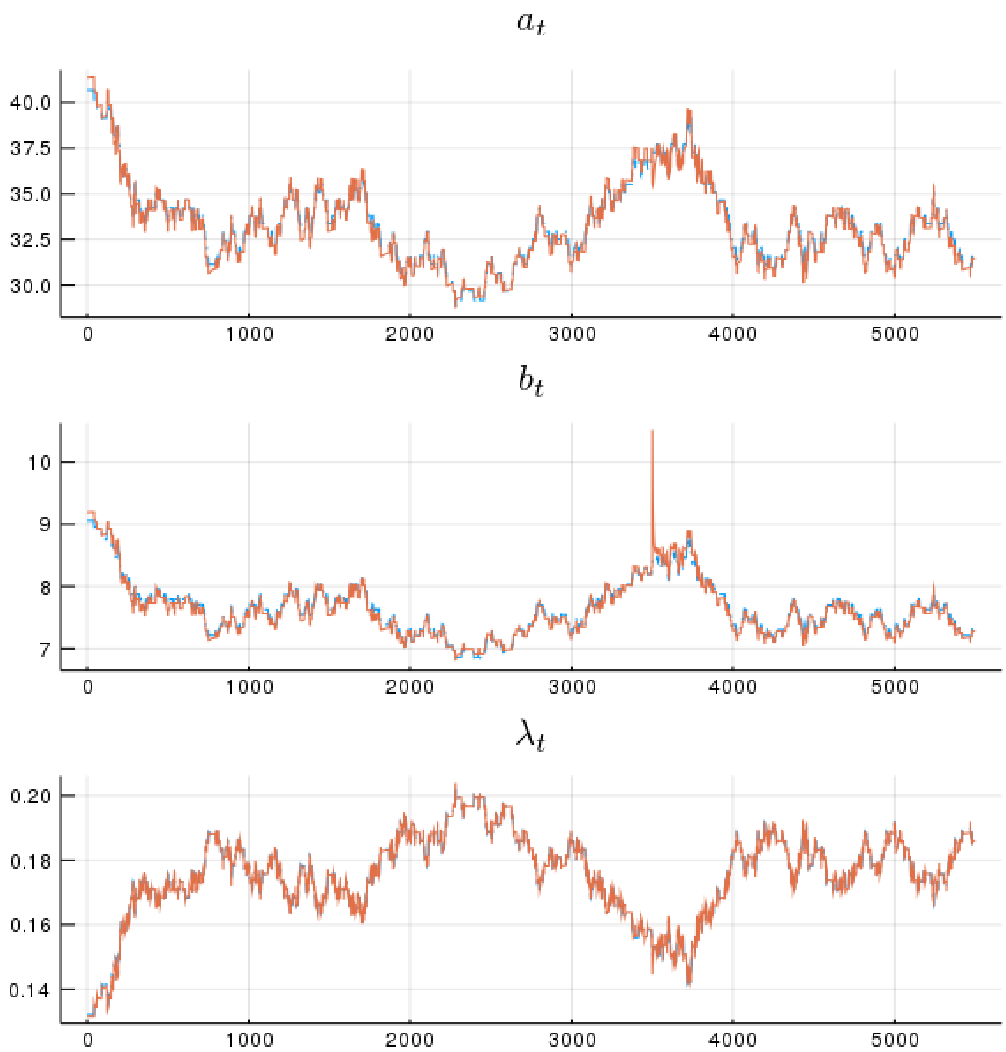
The forecasting rules for the shape parameter of the Beta density are given by:

$$a_t = 129.236 - 20.121\lambda_{t-1} - 90.718z_t$$

$$b_t = 18.227 - 21.445\lambda_{t-1} - 6.704z_t$$

While the list of potential predictors contains higher order terms and interaction terms, the LASSO selects only first order terms.

Figure 7: Accuracy of forecasting rules



Notes: This figure shows the series for the job meeting rate λ_t and for the shape parameters of the Beta distribution a_t and b_t approximating the distribution of vacancies across firm types. The solid orange lines represent the actual realization of the series. The blue lines are the series implied by the forecasting rules.

Table 12: Forecasting rule accuracy for λ_t

	Within-sample	Out-of-the-sample
max abs % error	13.35	13.28
min abs % error	5.65e-6	3.59e-6
mean abs % error	0.27	0.22
median abs % error	0.13	0.08

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted $\hat{\lambda}$ instead of the actual λ, calculated as $|\frac{\lambda-\hat{\lambda}}{\lambda}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

Table 13: Forecasting rule accuracy for a_t

	Within sample	Out of the sample
max abs % error	35.09	35.10
min abs % error	3.70e-5	1.00e-5
mean abs % error	0.96	0.76
median abs % error	0.69	0.59

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted $\hat{a}$ instead of the actual a, calculated as $|\frac{a-\hat{a}}{a}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

Table 14: Forecasting rule accuracy for b_t

	Within sample	Out of the sample
max abs % error	20.53	20.19
min abs % error	5.76e-6	2.61e-5
mean abs % error	0.87	0.75
median abs % error	0.65	0.56

Notes: This table shows the maximum, the minimum, the mean and the median absolute percentage error made when using the predicted $\hat{b}$ instead of the actual b, calculated as $|\frac{b-\hat{b}}{b}|$. The left column reports the prediction error in the sample used to calculate the forecasting rule; the right column shows the error made in a new sample, without re-estimating the forecasting rule.

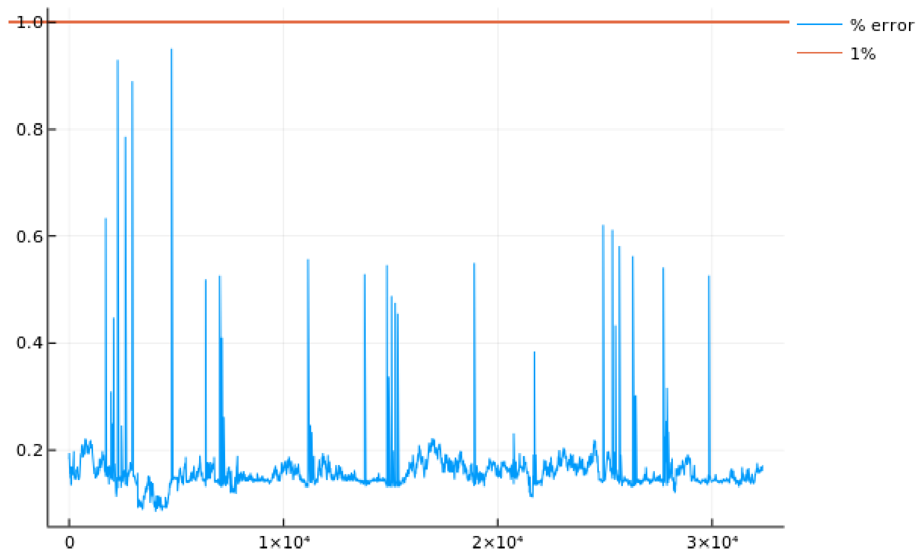
C.2 Accuracy of approximating $q_t(y)$ using a parametric function

I approximate the distribution of vacancies across types $q_t(y)$ using a parametric function $\tilde{q}(y|\boldsymbol{q_t})$. Because the support for y is $[0,1]$ and $q(y)$ is uni-modal, I use a Beta density characterized by two parameters a_t and b_t. To quantify the error made when approximating the endogenous distribution of vacancies by a parametric counterpart, I use the following measurement:

$$e_t = \frac{\int_0^1 c\left(|\tilde{q}(y|(\hat{a}_t, \hat{b}_t)) - q(y, \Gamma_t)|\right) dy}{\int c\left(q(y, \Gamma_t)\right) dy} \tag{34}$$

where $\hat{a}_t$ and $\hat{b}_t$ are the shape parameters of the Beta density, calculated using the forecasting rule and $c(v)$ is the cost of posting v vacancies. The variable e_t measures the cost of miss-allocated vacancies (the numerator), relative to the total cost of vacancy posting (the denominator). If agents are perfectly rational, $q(y, \Gamma_t) = \tilde{q}(y|(\hat{a}_t, \hat{b}_t))$ and the numerator is null.

Figure 8: Percentage error when approximating $q_t(y)$



Notes: This graph shows e_t, which is a unit-less measurement of cost miss-allocation implied by using $\tilde{q}(y|(\hat{a}_t, \hat{b}_t))$ to approximate $q(y, \Gamma_t)$. The variable e_t is defined by

$$e_t = \frac{\int_0^1 c\left(|\tilde{q}(y|(\hat{a}_t, \hat{b}_t)) - q(y, \Gamma_t)|\right) dy}{\int_0^1 c\left(q(y, \Gamma_t)\right) dy} \tag{35}$$

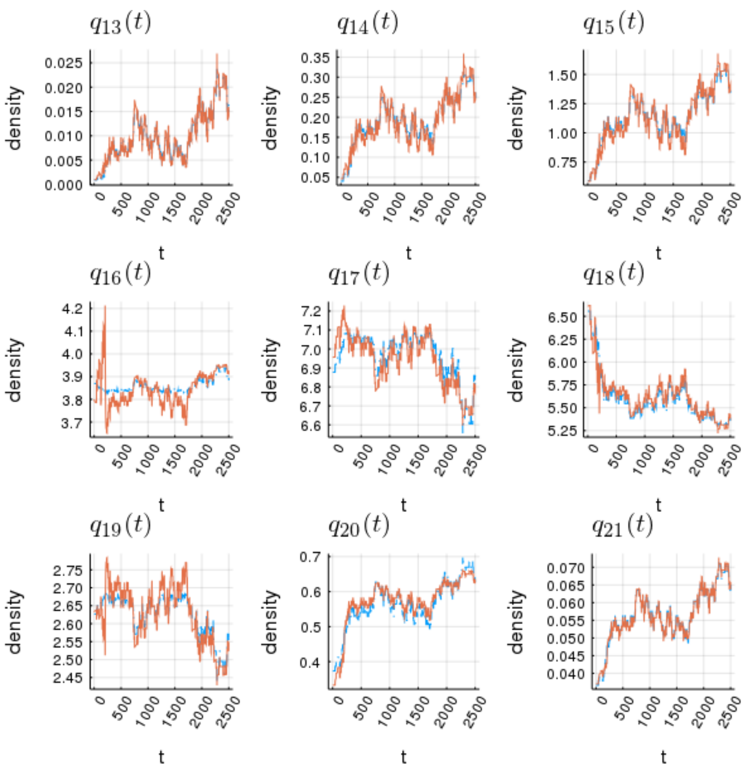
where $c(v)$ is the cost of posting v vacancies; $\hat{a}_t$ and $\hat{b}_t$ are the shape parameters of the Beta density implied by the forecasting rule. The maximum value for e_t is 0.950% and its median value is 0.151%.

C.3 Accuracy of approximating $q_t(y)$ using a histogram

One may dispense with the parametric assumption on the distribution of vacancies across firm types $q_t(y)$. Instead, as in Reiter (2009), one may use a histogram to approximate

$q_t(y)$. If the economy features N discrete firm types instead of a continuum, this approach amounts to assuming that agents forecast the number of vacancies posted by each firm type: $\boldsymbol{q_{t+1}} = (q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y)) = f_q(\Omega_t|\Theta_q))$. This approach is without loss of generality, because the y-dimension is already discretized when the model is numerically solved. Agents are endowed with a linear forecasting rule $\boldsymbol{q_{t+1}} = \Theta_q\Omega_t$, where $\boldsymbol{q_{t+1}}$ is a $N \times 1$ vector, Θ_q a $N \times k$ matrix containing the parameters for the forecasting rule, and Ω_t a $k \times 1$ matrix containing the information at time t relevant to predict $\boldsymbol{q_{t+1}}$. Ω_t contains $(z_t, z_{t-1}, \lambda_{t-1})$, their squares and interaction terms. As illustrated in Figure 9, the estimated forecasting rule is successful in predicting the distribution of vacancies across firm types.

Figure 9: Forecasting rules for $\boldsymbol{q_{t+1}}$



Notes: This figure shows selected components of the distribution of vacancies across firm types $\boldsymbol{q_{t+1}} = (q_{1;t+1}(y), q_{2;t+1}(y), ..., q_{N;t+1}(y))$. The orange lines are the actual realizations of the series and the blue lines are the values predicted by the forecasting rule $\boldsymbol{q_{t+1}} = \Theta_q\Omega_t$. Omitted components for $\boldsymbol{q_{t+1}}$ have a negligible value.

D BKM Algorithm

Throughout my exposition, I use the recursive formulation of a dynamic choice problem. However, the algorithm designed by Boppart et al. (2018) using the sequence form, is particularly attractive in my setting. The BKM algorithm uses the information contained in the perfect foresight path of the economy after a MIT shock. In general, a shooting algorithm has to be used to find this perfect foresight path, for which convergence properties are difficult to know a priori. In the present context, because the model is *semi-block recursive*[13], the algorithm of Boppart et al. (2018) can be easily used to find an approximation of the model. Semi-block recursivity ensures that BKM is well-behaved, because the path for $\lambda_t, q_t(y)$ can be solved without reference to $W_t(x, y)$. In practice, one could proceed as follows:

1. Solve $S(.)$ and $U(.)$ by value function iteration
2. Starting from the steady-state with no aggregate uncertainty at $t = 0$, generate a one-standard-deviation aggregate shock at time $t = 1$, which goes back to its steady-state value ($z = 1$) at $t = 2$.
3. Solve for the transition path of for $t = 2, ..., T$.
4. Solve $W_t(x, y, \lambda_t, q_y(t))$ by backward induction, starting from $W_t(x, y, \lambda_t, q_y(t)) = W_{SS}(x, y, \lambda_{SS}, q_{SS}(y))$

The perfect foresight path is obtained in a single step. If the model were not semi-recursive, finding the perfect foresight transition path would be much more complicated. The algorithm would have to be modified, with the addition of an outer loop, with no guarantee of convergence:

1. Assume a path for endogenous economic variables X_t, including $\{\lambda_t, q_y(t)\}$
2. Solve $S_t(.)$, $U_t(.)$ and $W_t(.)$ by backward induction, using the path previously assumed
3. Using $S_t(.)$, $U_t(.)$ and $W_t(.)$, simulate forward the path of economic variables, generating Y_t

[13] The employment problem is independent from wages. The wage problem depends on employment only through the job meeting rate and the distribution of vacancies.

4. If the distance between the paths X_t and Y_t is sufficiently small, stop
5. Otherwise, repeat the steps 1-4

The main advantage of using the BKM algorithm over the method presented in the full text, is that the parametric assumption for the distribution of vacancies is not required anymore.

E Flow equations for the joint distribution of wage and employment

Two methods are available to simulate the joint distribution of wages and employment status, denoted by $h_t(x,y,w)$. The first one is to simulate a panel with a sufficient number of agents. Then second one is to directly simulate the cross-sectional distribution $h_t(x,y,w)$. In this section, I derive the flow equations for the second approach. Let Ω_t denote the support for wages at period t. Given our assumption on the wage process, Ω_t contains the starting wages and the promotions wages offers within that period $\{\phi^0(x,y,\hat{\Gamma}_t),\phi^1(x,y,\hat{\Gamma}_t)\}_{(x,y)\in[0,1]^2}$. Ω_t also contains the wages inherited from past periods that were not altered by outside job offers or by intra-firm re-bargaining. The flow equation for the distribution of starting wages $h_t(x,y,\phi^0(x,y,\hat{\Gamma}_t))$ solves:

$$\begin{aligned}h_{t+1}(x,y,\phi^0(x,y,\hat{\Gamma}_t)) = {} & h_{t+}(x,y,\phi^0(x,y,\hat{\Gamma}_t)) \times \left(1-\int_0^1 s\lambda_t\frac{v_t(y')}{V_t}\mathbb{1}\{S(x,y',z_t)\geq S(x,y,z_t)\}dy'\right)\\ & + u_{t+}(x)\lambda_t\frac{v(y,\hat{\Gamma}_t)}{V(\hat{\Gamma}_t)}\mathbb{1}\{S(x,y,z_t)\geq 0\}\\ & + \int_{w\in\Omega_t\backslash\phi^0(x,y,\hat{\Gamma}_t)}\mathbb{1}\{W(x,y,w,\hat{\Gamma}_t)-U(x,z)<0\}(1-s\lambda_t+s_{x,y})h_{t+}(x,y,w)dw\end{aligned} \tag{36}$$

where $h_{t+}(x,y,\phi^0(x,y,\hat{\Gamma}_t)) \equiv (1-\delta)\mathbb{1}\{S(x,y,z)\geq 0\}h_t(x,y,\phi^0(x,y,\hat{\Gamma}_t))$ denotes the measure of workers with wage $\phi^0(x,y,\hat{\Gamma}_t)$ after endogenous and exogenous job destruction. The measure of employed workers meeting with a firm, which is not a threat to the current match, is denoted by $s_{x,y} \equiv \int_0^1 s\lambda_t\frac{v_t(y')}{V_t}\mathbb{1}\{S(x,y,z_t)\geq 0 > S(x,y',z_t)\}dy'$. The first line in (36) takes into account the outflow of workers poached by more productive firms. The second line considers the inflow of hiring from the pool of unemployed workers. The third line takes into consideration intra-firm re-bargaining. The flow equation for the measure at the promotion wage $h_t(x,y,\phi^1(x,y,\hat{\Gamma}_t))$ solves:

$$
\begin{aligned}
h_{t+1}(x,y,\phi^1(x,y,\hat{\Gamma}_t)) = h_{t+}(x,y,\phi^1(x,y,\hat{\Gamma}_t)) \times (1 - \int_0^1 s\lambda_t \frac{v_t(y')}{V_t} \mathbb{1}\{S(x,y',z_t) \geq S(x,y,z_t)\}dy') \\
+ \int_{w \in \Omega_t \setminus \phi^1(x,y,\hat{\Gamma}_t)} \mathbb{1}\{W(x,y,w,\hat{\Gamma}_t) - U(x,z) > S(x,y,z_t)\} \times \\
(1 - s\lambda_t + s_{x,y}) h_{t+}(x,y,w)dw
\end{aligned}
\tag{37}
$$

The first line in (37) takes into account the outflow of workers poached by more productive firms. The second and third lines take into consideration the measures of matches in which the firm had a credible threat to break the match. The expression for $h_{t+1}(x,y,\phi^1(x,y',\hat{\Gamma}_t))$ when $y' \neq y$ has to take into account workers poached by more productive firms, poaching from less productive firms and wage increases resulting from counter-offers:

$$
\begin{aligned}
h_{t+1}(x,y,\phi^1(x,y',\hat{\Gamma}_t)) = h_{t+}(x,y,\phi^1(x,y',\hat{\Gamma}_t)) \times (1 - \int_0^1 s\lambda_t \frac{v_t(y'')}{V_t} \mathbb{1}\{S(x,y'',z_t) \geq S(x,y,z_t)\}dy'') \\
+ \int_0^1 s\lambda_t h_{t+}(x,y') \frac{v_t(y)}{V_t} \mathbb{1}\{S(x,y,z_t) \geq S(x,y',z_t)\}dy \\
+ \int_{w \in \Omega_t \setminus \phi^1(x,y',\hat{\Gamma}_t)} \int_0^1 s\lambda_t h_{t+}(x,y,w) \mathbb{1}\{\phi^1(x,y',\hat{\Gamma}_t) \geq w\} \frac{v_t(y')}{V_t} \times \\
\mathbb{1}\{S(x,y,z_t) > S(x,y',z_t) \geq 0\}dy'dw
\end{aligned}
\tag{38}
$$

For wages that are not in the set of starting or promotion wages denoted by $\{\phi^0(x,y,\hat{\Gamma}_t), \phi^1(x,y,\hat{\Gamma}_t)\}_{(x,y)}$, the flow equation takes into account workers (i) surviving both endogenous and exogenous job destruction (ii) workers with no on-the-job meeting or meeting (or choosing not to disclose unsuccessful ones) (iii) workers with no intra-firm re-bargaining:

$$
h_{t+1}(x,y,w) = h_{t+}(x,y,w)(1 - s\lambda_t + s_{x,y}) \times \mathbb{1}\{0 \leq W(x,y,w,\hat{\Gamma}_t) - U(x,z) < S(x,y,z_t)\}
\tag{39}
$$

F Identification

Table 15: Jacobian of $f : \boldsymbol{p} \rightarrow \boldsymbol{m}$

Moment	α	δ	s	β_1	β_2	c_0	c_1	p_1	p_2	p_3	p_4	p_5	p_6	b_0	b_1	ϕ
U	-0.07	1.76	0.24	-0.05	0.00	0.19	0.00	-0.03	0.00	-0.01	0.00	0.00	0.00	0.05	0.05	0.00
U_{5+}	-0.02	0.24	0.06	-0.04	0.00	0.05	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00
U_{15+}	0.00	0.00	0.00	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
U_{27+}	0.00	0.00	0.00	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\frac{U}{V}$	4.78	-114.66	-17.38	4.07	-0.33	-31.70	3.43	4.37	0.71	1.07	0.15	0.08	0.05	-7.61	-7.66	1.07
$U2E$	-0.35	17.70	1.28	0.89	-0.03	1.00	-0.04	-0.15	-0.02	-0.03	0.00	0.00	0.00	0.23	0.24	-0.08
$J2J$	0.06	1.08	-0.04	0.02	0.00	-0.17	0.02	0.02	0.00	0.00	0.00	0.00	0.00	-0.04	-0.04	0.00
$E2U$	-0.05	1.59	0.19	-0.01	0.00	0.15	0.00	-0.02	0.00	-0.01	0.00	0.00	0.00	0.04	0.04	0.00
std_U	0.00	3.18	-0.43	0.17	-0.01	-0.01	0.16	-0.10	-0.01	-0.01	0.00	0.02	0.01	0.31	0.24	1.29
$std_V U$	0.00	3.62	-0.41	0.17	-0.01	0.01	-0.04	-0.11	-0.01	-0.01	0.00	0.02	0.01	0.65	0.49	2.97
std_{E2U}	0.29	-1.06	-1.53	0.11	-0.01	-0.84	0.19	0.02	0.01	0.02	0.00	0.02	0.01	0.14	0.06	1.46
std_{U2E}	0.25	-3.56	-0.95	-0.05	0.00	-0.70	0.03	0.10	0.02	0.02	0.00	0.00	0.00	-0.14	-0.15	0.18
std_V	-0.01	0.45	0.01	-0.01	0.00	0.02	-0.20	-0.01	0.00	-0.01	0.00	0.00	0.00	0.34	0.25	1.68
std_U	0.00	3.18	-0.43	0.17	-0.01	-0.01	0.16	-0.10	-0.01	-0.01	0.00	0.02	0.01	0.31	0.24	1.29
std_{U5+}	-0.46	8.75	1.35	0.19	-0.01	1.33	0.10	-0.29	-0.04	-0.05	-0.01	0.02	0.01	0.60	0.54	1.03
std_{U15+}	-0.04	0.52	0.14	0.00	0.00	0.12	0.00	-0.02	0.00	0.00	0.00	0.00	0.00	0.04	0.04	0.04
std_{U27+}	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
std_{GDP}	-0.01	0.28	0.02	0.00	0.00	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.47
std_{J2J}	-0.04	2.11	-0.51	0.01	0.00	0.12	0.03	-0.06	-0.01	-0.02	0.00	0.00	0.00	0.26	0.20	1.04
$corr(E2U, GDP)$	-0.02	-0.06	0.09	-0.01	0.00	0.04	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.27
$(corr(V, GDP)$	0.00	0.13	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16
$corr(U2E, GDP)$	-4.88	72.69	18.10	1.11	0.03	13.96	-0.48	-2.06	-0.32	-0.47	-0.06	-0.01	-0.01	3.19	3.27	-1.75
$corr(U, GDP)$	0.00	-0.20	-0.01	0.00	0.00	-0.01	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	-0.15
$corr(U, V)$	0.00	0.02	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$corr(U2E, J2J)$	-5.12	76.49	18.88	1.16	0.03	14.64	-0.41	-2.22	-0.35	-0.56	-0.07	-0.03	-0.02	3.50	3.55	-1.14
$autocorr(VA)$	-0.01	0.21	0.01	0.00	0.00	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.34
$elasticity_P 10$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$elasticity_P 20$	-1.33	-82.78	-13.40	-0.08	-0.02	3.81	-0.80	0.34	0.05	0.88	-0.03	0.47	0.13	-3.68	-3.64	-0.38
$elasticity_P 30$	-29.27	-1815.16	-293.20	-4.28	0.17	83.70	-16.58	7.66	1.01	14.81	-0.66	8.44	2.25	-82.26	-80.81	-26.38
$elasticity\_P40$	120.59	-457.84	1227.69	29.82	-3.16	-344.87	45.59	66.54	11.99	14.76	0.10	0.67	-0.10	-10.41	-8.35	-80.54
$elasticity_P 50$	61.92	47.30	635.25	14.84	-1.54	-177.09	23.58	30.80	5.60	6.63	0.19	0.57	0.10	6.09	6.97	-16.92
$elasticity_P 60$	231.49	172.99	2364.13	54.36	-5.51	-662.01	99.83	121.41	20.38	23.37	0.23	-2.10	-0.36	39.60	41.01	-40.17
$elasticity_P 70$	149.67	155.89	1530.98	35.27	-3.60	-428.03	63.61	77.49	12.98	12.10	0.27	-0.95	-0.17	26.00	26.90	-22.13
$elasticity_P 80$	50.52	85.40	517.90	11.95	-1.22	-144.49	21.52	24.26	4.16	1.17	0.11	-1.12	-0.28	9.65	10.31	-11.23
$elasticity_P 90$	-26.62	13.48	-288.15	-7.20	0.88	76.14	-5.84	-9.78	-2.11	-3.41	-0.08	-0.44	-0.35	-10.42	-9.92	-85.40

G Inference

Because the likelihood function is untractable, I use the Simulated Method of Moments (SMM) to estimate parameter values (see for instance Duffie and Singleton (1990) and Gourieroux et al. (1993)). The SMM estimates are the one minimizing a weighted difference of simulated moments from their empirical counterparts:

$$\hat{\boldsymbol{\theta}}_{SMM} = \arg\min_{\boldsymbol{\theta}} \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\boldsymbol{\theta})\right) W \left(\hat{\boldsymbol{m}} - \boldsymbol{m}^S(\boldsymbol{\theta})\right)' \tag{40}$$

where $\hat{\boldsymbol{m}}$ is a vector of empirical moments and $\boldsymbol{m}^S(\boldsymbol{\theta})$ a vector of the same moments, calculated using simulated data. More specifically, the l element of the vector $\hat{\boldsymbol{m}}$ is calculated as $m_l \equiv \frac{1}{T}\sum_{t=1}^{T} f_{l,t}^*$, where $f_{l;t}^* \equiv f_l(Y_t, Y_{t-1}, ..., Y_{t-k+1})$ with f_l a function mapping the finite l-history of state information $\{Y_t, Y_{t-1}, ..., Y_{t-k+1}\}$ to $\mathbb{R}$. I choose the weighting matrix W to be diagonal, with values representing subjective weight of the moments I deem more important to match.

In this paper, my goal is to match unconditional moments from a time series. In this context, under some regularity conditions[14], the SMM estimate is asymptotically normally distributed with asymptotic variance $(1+\tau)$ times that of the GMM estimator, where $\tau = \frac{T}{S}$, with T the sample size and S the length of the simulated sample. As the size of the simulated sample increases relative to the actual sample size, the efficiency loss due to using simulated data rather than real ones becomes negligible. Note that $\hat{\theta}_{SMM}$ is a function of both τ and the weighting matrix W. The SMM estimator has the following formula for the asymptotic variance:

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{SMM} - \boldsymbol{\theta}_0\right) \overset{A}{\sim} \mathcal{N}(0, (1+\tau)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}) \tag{41}$$

whith $\lim_{T\to\infty} \frac{T}{S(T)} = \tau$

$$\Sigma_0 = \sum_{j=-\infty}^{+\infty} \mathbb{E}\left([\boldsymbol{f}_t^* - \mathbb{E}(\boldsymbol{f}_t^*)][\boldsymbol{f}_{t-j}^* - \mathbb{E}(\boldsymbol{f}_{t-j}^*)]'\right)$$

$$\Sigma_1 = D'WD$$

$$\Sigma_2 = D'W\Sigma_0 WD$$

[14] see Gourieroux et al. (1993), page 31

$$D = \mathbb{E}_0 \left(\frac{\partial \boldsymbol{m}^S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right)$$

In practice, I approximate D numerically using a finite difference scheme. To calculate Σ_0, I use a HAC estimator on simulated data rather than on real data. Given that the convergence rate of spectral estimators is low and that I control the length of the simulated sample, this potentially increases the accuracy of my estimate, as discussed in Duffie and Singleton (1990). The standard error for the estimate l is then calculated by taking the square root of the (l,l) element of the estimate for the asymptotic matrix, multiplied by $\frac{1}{T}$:

$$SE_l = \left(\frac{1}{T} \left((1+\tau)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1} \right)_{ll} \right)^{1/2}$$