Accounting for idiosyncratic wage risk over the business cycle*

Alisdair McKay† and Tamás Papp‡

February 15, 2012

Abstract

We demonstrate that wage volatility, measured as the cross-sectional variance of wage changes in PSID data, is positively correlated with the unemployment rate, and thus it is countercyclical. We quantify this relationship by estimating the regression coefficient of wage volatility on the national unemployment rate in a multilevel Bayesian model, then decompose this coefficient into three main factors. During a recession, wage volatility increases substantially among those workers experiencing spells of unemployment: the cyclical changes in the variance within this group explain about 55% of the cyclical variation in wage volatility. The variance within the group not experiencing unemployment explains 18%. Finally, an increase in the fraction of workers experiencing unemployment explains 25%.

We show that a parsimoniously calibrated search-and-matching model of the labor market with on-the-job search gives a good account of the cyclical variation in idiosyncratic wage risk among those experiencing unemployment and of the composition effect over the business cycle. As these components generate around 80% of the total, we argue that such a model gives a good account of cyclical variation in idiosyncratic wage risk.

JEL classification: C11, E24, E32, J64

*This research is supported by the Austrian Central Bank Grant No: 14076. We thank Christian Haefke, Nobuhiro Kiyotaki, Per Krusell, Monika Merz, Michael Reiter, Christopher Sims, participants in seminars at the Boston University, the Federal Reserve Banks of Richmond and San Francisco, CERGE-EI, CESifo, NORMAC, and the Vienna Macro Breakfast. Any remaining errors are our own.

†Boston University. amckay@bu.edu

‡Institute for Advanced Studies, Vienna. tpapp@ihs.ac.at

1 Introduction

In this paper we (1) document the countercyclical variation in the variance of wage shocks in PSID data, (2) decompose this countercyclicality using unemployment history to assess the quantitative importance of various components, demonstrating that most (55%) of the countercyclicality is explained by agents who undergo an unemployment spell, with the composition effect (the cyclicality of unemployment incidence) explaining about 25%, (3) demonstrate that a parsimoniously calibrated search-and-matching model is able to quantitatively explain these sources of countercyclical variation in the variance of wage shocks.

The motivation for our work comes from the literature that models aggregate fluctuations with heterogeneous households, in which Krusell and Smith (1998) is the classic reference. The connection between aggregate and idiosyncratic risk is an important open question in this literature. Depending on the nature of the relationship between aggregate events and idiosyncratic shocks, aggregate stabilization policies may be able to affect individual households by changing the nature of the idiosyncratic risks that those households face. It is therefore necessary to understand the mechanisms that link aggregate events, policies, and shocks to the individual outcomes of households in order to fully understand the welfare consequences of alternative policies.

Existing research has taken two approaches to modeling the link between aggregate shocks and idiosyncratic shocks. One approach, for example Krusell and Smith (1998), takes the connection between the aggregate state of the economy and the variance of idiosyncratic shocks to be a statistical fact that is fed into the model as an exogenous feature of the income process. Under this approach, authors interested in counterfactual experiments, usually the elimination of business cycles, have relied on assumptions such as the integration principle, which posits that the distribution of idiosyncratic risks in the counterfactual should be constructed by integrating the distribution of idiosyncratic shocks across aggregate states.[1] The other approach has been to incorporate structural models of the labor market in which the nature of income shocks is determined endogenously.[2]

[1] The following papers have implemented the integration principle or related approaches to analyze the welfare cost of business cycles: Imrohoruglu (1989), Krusell and Smith (1999), Storesletten et al. (2001), Krebs (2003, 2007) Krusell et al. (2009). Atkeson and Phelan (1994) argue that idiocyncratic risk could be completely unchanged in the absence of business cycles if the business cycle is entirely the result of the correlation of shocks across individuals.

[2] Gomes et al. (2001) and Beaudry and Pages (2001) have analyzed the welfare cost of business cycles using models in which incomes are determined endogenously. We will discuss the work of Gomes et al. and how it relates to our

A natural question is to ask which of these structural models of the labor market generate wage and earnings dynamics that are consistent with the observed data on individual wages and earnings.

Our chief object of interest is the cross-sectional variance of year-to-year changes in log wages, which we call wage volatility. Our focus on this feature of the wage process is motivated by several factors. We focus on wages because they are, in a sense, more fundamental than earnings, which reflect wages as well as labor supply decisions (see Heathcote et al., 2010). We study the dispersion in growth rates, as opposed to levels, because they are a plausible indicator of the size of unanticipated and uninsurable shocks to individuals. Finally, we use the variance to summarize the dispersion because it allows us to decompose the overall dispersion into different sources. There are other features of the wage process beyond the innovation variance that are also of importance to individual households, notably the persistence of the shocks, which we do not consider here. A fully satisfactory structural model of the wage process should generate realistic predictions on these dimensions as well.

We begin by documenting the extent to which the variance of wage shocks varies over the business cycle. To do so, we use a multilevel statistical model that relates the variance of wage shocks to the national unemployment rate. We estimate this model using data on men's wages from the PSID and the national unemployment rate reported by the BLS. The main result of section 2 is the demonstration of a strong comovement between wage volatility and the unemployment rate. We also find evidence of countercyclical volatility in annual earnings. It is not surprising that there is countercyclical volatility in earnings since the incidence and variability of unemployment shocks increase during a recession.

After documenting the magnitude of countercyclical wage volatility, the next step in our analysis is to investigate its sources. Standard frictional models of the labor market (Mortensen and Pissarides, 1994; Burdett and Mortensen, 1998) suggest that workers redraw from the wage distribution after experiencing an unemployment spell, which should, by construction, increase cross-sectional

empirical facts. Beaudry and Pages demonstrate that aggregate wage volatility may understate the wage risk faced by individuals if employers provide insurance for the wages of workers in continuing matches. This model of the labor market, however, has the implication that aggregate stabilization eliminates all idiosyncratic risk to workers. Recently, a number of authors have studied versions of the Diamond-Mortensen-Pissarides search and matching model with risk averse workers and imperfect consumption insurance.See Costain and Reiter (2005, 2007); Shao and Silos (2007); Rudanko (2009, 2011); Nakajima (2010); Bils et al. (2011); Krusell et al. (2010); Jung and Kuester (2011).

volatility when unemployment increases. Consequently, we partition the sample into groups according to whether an individual has experienced unemployment in the previous two years (i.e. the years over which we take the the first-difference in wages). We then perform a variance decomposition exercise that relates the total variance of wage growth (i.e. wage volatility) to the within group variances, the within group means and the group sizes and we explore how these components vary with the unemployment rate.

We show that the comovement of wage volatility and the unemployment rate is in part driven by a composition effect. The composition effect arises because more people experience unemployment in a recession and unemployment is associated with higher wage volatility at all times. Therefore, during a recession, the composition of the sample shifts towards the group with the larger variance. We find that this effect explains about 25% of the relationship between wage volatility and the unemployment rate. The other important factors are the cyclical changes in the within group variances. In particular, the variance of wage shocks among those experiencing unemployment increases strongly with the unemployment rate. This effect can explain about 55% of the cyclical movements in overall wage volatility, which is notable because these observations make up only around one quarter of our sample. Increasing volatility among those not experiencing unemployment explains most of the remainder, about 18%. These results are estimated with little posterior uncertainty, and we demonstrate that this is an advantage of using a multilevel model.

In sections 4, 5 and 6, we turn our attention to the economic mechanisms generating the data. Search-and-matching models provide natural links between aggregate labor market conditions and the experiences of individual workers. We ask whether models in this class can explain the cyclical variation in wage volatility that we document in the data. We consider a prototypical search-and-matching model with on-the-job search and ask whether the observed variation in the job-offer and separation rates can explain the wage dynamics that we observe. The model does a good job of explaining the countercyclical variation in the variance of wage innovations among those experiencing unemployment and the composition effect. Together, these components generate 80% of the total effect that we have identified in the data.

The key mechanism at work in the model is that the job-separation rate increases and the job-offer rate decreases in a recession, both of which lead an unemployed worker to reduce his reservation wage. A lower reservation wage means that there is a larger support of the wage distribution and this leads to more volatile wages as workers are climbing up and falling off a taller wage ladder. We explore this mechanism analytically and with a quantitative experiment.

In section 4, we set up our partial equilibrium model, then focus on the elasticity of the reservation wage with respect to changes in the job-offer and separation rates in steady state. Studying the steady state is useful because the model economy converges to the steady state quickly and so changes to the steady state give a good approximation to the full dynamic behavior of the economy. This analysis allows us to identify the parameters of the model and the features of the distribution of match-specific productivity levels that determine the sensitivity of the reservation wage to the job-finding and separation rates.

Our analytical results guide our calibration of the model in section 5. We then proceed in Section 6 to simulate the distribution of wage growth rates. In our simulation exercise, we feed job-offer and separation rates into the model in such a way that the model replicates the observed job-finding and separation rates reported by Shimer (2007). This simulation exercise reveals that the model generates the cyclical variation in the variance of growth rates among those workers experiencing unemployment and the composition effect, much as we observe in the data. Section 7 concludes.

Related literature Our results in section 2 are related to work by Storesletten et al. (2004) who estimate an income process that allows the variance of shocks to differ between expansions and contractions. As our methodology differs from theirs, it is worth considering its advantages and disadvantages. Storesletten et al. point out that if the variance of persistent income shocks increases during a recession then cohorts who have lived through more recessions will have a larger variance of income levels. They then use variation in macroeconomic experiences across cohorts to estimate the extent to which the variance of income shocks increases in a recession. The key to their approach is that they recognize that data on incomes from 1967 onwards contain information about income shocks that occurred since 1930. The difficulty of this approach for our purposes is that we only have covariates, such as data on unemployment spells, since 1967 so we are unable to investigate the sources of income shocks before the start of our sample. In comparison to

Storesletten et al., we use information on a more limited set of of aggregate fluctuations. Despite the fact that we only use information between 1967 and 1992, we identify a clear and significant countercyclical pattern in wage volatility. While the focus of our paper is on wages, we note in section 2.3 that when we apply our methods to data on incomes the resulting estimates are in line with those of Storesletten et al..

Similarly, our work is related to that of Meghir and Pistaferri (2004) in that we are also interested in changes in the variance of idiosyncratic shocks over time. Their work is focused on identifying ARCH effects in the variance of idiosyncratic shocks rather than on connection between those variances and the aggregate labor market. Moreover, the structural mechanisms underlying the income process that Meghir and Pistaferri estimate is something of an open question.

Variance decomposition techniques have previously been used to understand secular trends in inequality. For example, Lemieux (2006) decomposes the trend in the variance of hourly wages from the CPS and shows that the shift towards an older, more educated workforce can explain a third to a half of the increase in residual wage inequality between 1973 and 2003.[3] Our approach differs from previous work because we focus on cyclical fluctuations as opposed to trends.

2 Wage Volatility and Unemployment

In this section we introduce our multilevel modeling approach and estimate the a linear relationship between the volatility of wages and the unemployment rate. The data we use are from the Panel Study of Income Dynamics (PSID) covering income in years 1967 to 1992.[4] We measure wages as the ratio of annual labor income to annual hours worked and deflate to 1967 dollars using the

[3]The sociology literature has also explored trends in income inequality and recently Western and Bloome (2009) have shown how to construct standard errors for Lemieux's decomposition using Bayesian methods.

[4]Two considerations influence our choice of years to include in our sample. First, our object of interest is the first-difference of log annual wages for which we need data on wages in consecutive years. Therefore we cannot make use of PSID data after 1996 (survey year 1997) after which the PSID switches to a biannual frequency. Second, there is a structural break in PSID wage volatility around 1993, which has also been documented by Heathcote et al. (2010). The timing of this break coincides with the switch to a computer-based survey methodology although it is not clear whether this break represents an actual change in the data generating process or if it is an artifact of the methodological change. In our analysis, we have found that our findings survive if we model this break as a level shift in the wage volatility process. We choose, however, to end our sample in 1992 for the sake of simplicity and ease of exposition. Finally, the switch to the new survey methodology began in survey year 1993 and was completed in survey year 1994 so our 1992 data have some elements of the new methodology. The inclusion of 1992 does not exert a strong influence on our results.

CPI-Research price index. We restrict the sample to male heads who were between the ages of 25 and 60 and worked at least 320 hours per year. Students, business owners, and self-employed individuals are excluded from the analysis. We focus on the first difference of log wages across years so individuals must be present for two consecutive years in order to be included in our sample. Appendix A provides further details about our sample.

2.1 Multilevel Model

Our statistical model consists of three equations. For an individual i in year t, we model the innovation in log wages as

$$\mathrm{d}w_{i,t} \sim \mathrm{N}\left(X_{i,t}\theta + \alpha_t, \sigma_t^2\right) \tag{1}$$

$$\alpha_t \sim \mathrm{N}\left(Z_t\kappa, {\varsigma_\alpha}^2\right) \tag{2}$$

$$\sigma_t^2 \sim \mathrm{N}\left(Z_t\eta, {\varsigma_{\sigma^2}}^2\right), \quad \sigma_t^2 > 0. \tag{3}$$

The first line states that the change in an individual's log wage, $\mathrm{d}w$, is normally distributed. The mean of this distribution depends on the individual's demographic characteristics such as age and education, which are placed in the vector $X_{i,t}$. We assume that the coefficients on these demographic characteristics are common across years. In addition, the mean change in wages varies over time with the α_t term. Finally, we allow the variance of the innovation in wages to vary over time as captured by the σ_t^2 term.

The second and third lines show the multiple levels of our model as we impose structure on the parameters of the wage growth distribution and assume that they are drawn from their own, estimated distributions.[5] The α_t terms are drawn from a normal distribution, the mean of which is linearly related to aggregate variables, Z_t. We use the national unemployment rate as our measure of aggregate conditions and this, along with a constant, makes up the vector Z_t. Similarly, equation (3) relates the variance of wage growth to aggregate conditions. As in equation (2), the mean of this distribution is linearly related to the national unemployment. Our main object of interest is the second element of η, that is the coefficient on the unemployment rate in equation (3), which

[5] See Gelman et al. (2004) and Gelman and Hill (2007) for a discussion of multilevel models.

we call η_{unemp}. As the variance must be positive, we draw from a truncated normal distribution — in practice, however, the mass of the distribution below zero is so small that this truncation is practically irrelevant.

By estimating all of the parameters of the model jointly, the posterior uncertainty about η reflects the uncertainty about σ_t^2. By contrast, one can imagine a two-stage estimation procedure in which one first estimates equation (1) and then computes σ_t^2 from the residuals of this first-stage regression and uses these estimates as "data" in estimating equation (3). The difficulty with this two-stage approach is that the standard errors in the second-stage regression do not reflect the fact that σ_t^2 is itself an estimate. The multilevel model avoids this difficulty by estimating both stages at once.

Another advantage of the multilevel model is that it provides sharper estimates of σ_t^2 than one would obtain from the first-stage regression. The reason is that the multilevel model is able to pool information across years if the data suggest that $\mathrm{d}w_{i,t}$ in those years are generated by a similar process. Consider the meaning of the parameter ς_{σ^2}. If the unemployment rate were the same in years t and s and $\varsigma_{\sigma^2} \approx 0$, then from equation (3) it follows that $\sigma_t^2 \approx \sigma_s^2$, which is to say that the variance of wage changes in years t and s is the same. If these variances are equal, we can estimate them more precisely by pooling the data from years t and s to estimate the single variance that applies to both years. Alternatively, if ς_{σ^2} is very large, the implication is that even if the unemployment rate is the same in years t and s we have no reason to think that σ_t^2 and σ_s^2 should be related. Therefore, we should estimate σ_t^2 and σ_s^2 separately without pooling the data. In between these extremes, the model can partially pool the data across years by down-weighting the data from year s when estimating σ_t^2.

The amount of pooling that actually occurs in estimating σ_t^2 depends on the value of ς_{σ^2} which is itself estimated jointly with the other parameters of the model. This is possible because the likelihood depends on the parameters of both levels of the model. If the data do not call for pooling, low values for ς_{σ^2} have low likelihood because the data require (relatively) large errors in equation (3) or require that equation (1) be fit with similar variances, which is at odds with the data. Conversely, if the data call for pooling, high values of ς_{σ^2} have low likelihood because the errors in equation (3) are small so the likelihood can be raised by reducing ς_{σ^2}. By following this logic, the multilevel model is able to pool data across years to the extent called for by the data. In section 3.2.1 we discuss exactly how much sharper our estimates are as a result of this partial pooling.

By including the unemployment rate in equation (3), we allow the model to attribute some of the differences in σ_t^2 across years to changes in the unemployment rate. If the unemployment rate explains some of the variation in σ_t^2, our statistical procedure automatically tightens our estimate of ς_{σ^2}. Since equation (3) is our prior for estimating σ_t^2, a lower value of ς_{σ^2} implies a sharper prior on σ_t^2. This prior, which is itself estimated from data across years, is the mechanism through which the model is able to use information from other years to inform the estimate of σ_t^2. When the prior becomes more precise, it has a larger effect on the estimate of σ_t^2 and more information is pooled across years. So if equation (3) fits better, ς_{σ^2} falls and more information is pooled. In effect, the inclusion of predictors, such as the unemployment rate, in equation (3), allows the model to identify dimensions on which we expect σ_t^2 to differ between years and therefore pool information more effectively.

2.1.1 Prior Distributions

We need to specify prior distributions for θ, κ, η, ς_α and ς_{σ^2}. Since the first three are regression coefficients, it is natural to use the non-informative (reference) priors

$$p(\theta) \propto 1 \tag{4}$$

$$p(\kappa) \propto 1 \tag{5}$$

$$p(\eta) \propto 1 \tag{6}$$

There is sufficient sample size at each level of the model to make the posterior distribution proper.

In the context of an ordinary (single-level) linear regression, the usual choice for a non-informative reference prior for the variance ς^2 of the error term is $p(\varsigma^2) \propto \varsigma^{-2}$. In the context of multilevel models, however, Gelman (2006) demonstrates that this prior places infinite mass near $\varsigma^2 = 0$, resulting in an improper posterior distribution, then suggests that weakly informative priors are used instead, highlighting the advantages of the conditionally conjugate folded-noncentral-t family.

Following the recommendation of Gelman (2006) and Polson and Scott (2011), we use a special case of this family, the half-Cauchy priors

$$p(\varsigma_\alpha) \propto ({\varsigma_\alpha}^2 + s_\alpha^2)^{-1} \tag{7}$$

$$p(\varsigma_{\sigma^2}) \propto ({\varsigma_{\sigma^2}}^2 + s_{\sigma^2}^2)^{-1}. \tag{8}$$

with $s_\alpha = s_{\sigma^2} = 1$.[6]

2.1.2 Estimation

We use a Gibbs sampler to draw from the posterior distribution of the parameters of the model. We partition the parameter space into blocks corresponding to (θ, α), κ, ς_α, σ^2, η, and ς_α and sample each block in turn. Many of the sampling steps reduce to drawing from an ordinary linear regression or conjugate distributions, otherwise we use the slice sampler of Neal (2003). Appendix B contains details on the estimation methodology.

2.2 Results

	mean	5%	25%	50%	75%	95%
θ_{age}	−0.0014	−0.0016	−0.0015	−0.0013	−0.0012	−0.0011
θ_{edu}	0.0028	0.0019	0.0024	0.0028	0.0032	0.0038
κ_{const}	0.0258	0.0202	0.0236	0.0257	0.0279	0.0313
κ_{unemp}	−0.6402	−1.0127	−0.7858	−0.6368	−0.4904	−0.2911
ς_α	0.0145	0.0101	0.0123	0.0141	0.0163	0.0201
η_{const}	0.0771	0.0741	0.0758	0.0771	0.0784	0.0802
η_{unemp}	0.4255	0.2257	0.3428	0.4244	0.5077	0.6311
ς_{σ^2}	0.0091	0.0068	0.0079	0.0089	0.0099	0.0119

Table 1: Posterior means and quantiles for model parameters.

[6] In an earlier draft of the paper we used uniform noninformative priors of the form $p(\varsigma) \propto 1$, which yields a proper posterior with enough (> 3) groups — since we have 25 years, this requirement is satisfied. However, Polson and Scott (2011) show that the half-Cauchy prior also has desirable frequentist properties, and also suggest that

> The half-Cauchy occupies a sensible 'middle ground' within this class [of hypergeometric inverted-beta priors]: it performs very well near the origin, but does not lead to drastic compromises in other parts of the parameter space.

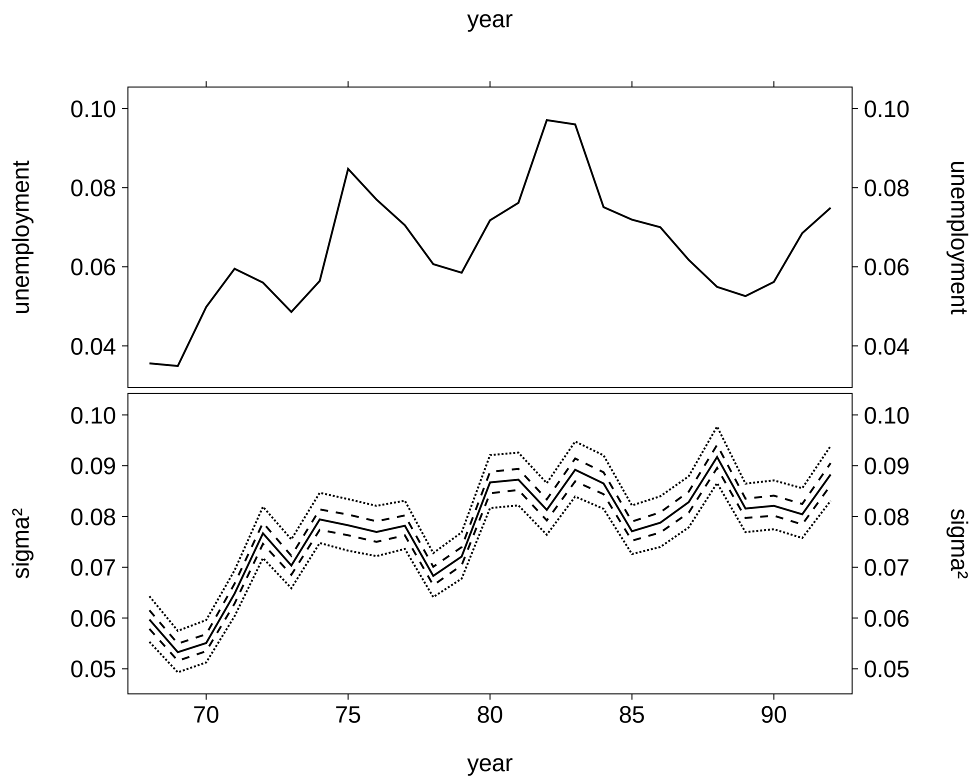

Figure 1: The top panel shows the unemployment rate, the bottom panel shows the posterior quantiles of σ_t^2. The solid line indicates the median, the dashed lines the 25% and 75% quantiles, and the dotted lines the 5% and 95% quantiles from the posterior.

Table 1 shows the posterior means and quantiles of the parameters of our model. The posterior mean and median of η_{unemp} are both 0.42 and the 90% error band extends from 0.22 to about 0.62. η_{unemp} can be thought of as the slope of a least squares regression of σ_t^2 on the unemployment rate. As such, a positive coefficient implies a positive co-movement. We also note that the sign of η_{unemp} is almost unambigously positive: less than 0.2% of the posterior draws are below 0.

To highlight the positive correlation between these two series, Figure 1 shows the posterior distribution of σ_t^2 along with the unemployment rate versus time, and Figure 2 shows a scatter plot of the posterior median of the σ_t^2 against the unemployment rate. The vertical bars in the figure are 90% error bars for σ_t^2. The figure also shows posterior draws of the regression line from equation (3). The logic of the multilevel model is that we parameterize the prior distribution on σ_t^2 and equation (3) represents the parameterization of this prior. The difference between the points

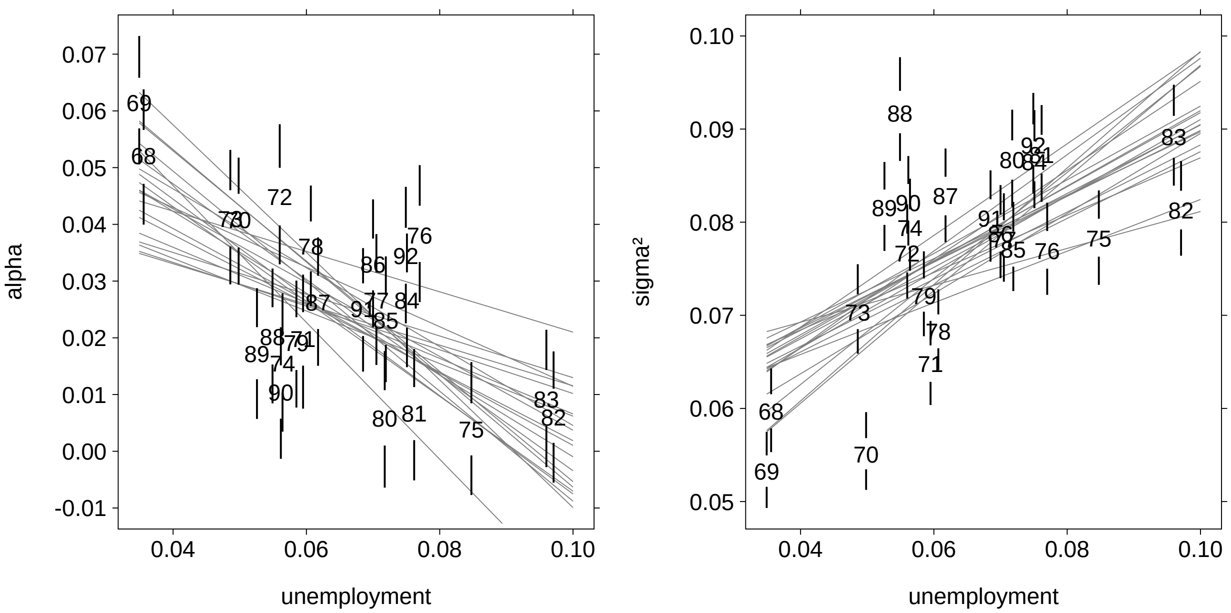

Figure 2: Scatter plot of posterior distribution of α_t (left panel) and σ_t^2 (right panel) versus the unemployment rate in year t. Year labels are placed at the posterior medians, vertical bars extend from the 25% to 5% quantiles and from the 75% to the 95% quantiles. The regression lines are random draws from the posterior distribution.

in the scatter plot and the sample of regression lines is that the regression lines show the variation in the prior on σ_t^2 and the scatter plot shows the estimated values of σ_t^2, which are a compromise between the data and the prior.

Even though we are primarily interested in η_{unemp}, the posterior distribution of the other parameters is in line with our expectations and the literature : κ_{unemp} is estimated to be negative, which implies that real wage growth is counter-cyclical. In the first-level of the model, the demographic effects show that wage growth declines with age and increases with education.

2.3 Relation to Previous Work

From the estimated relationship between the unemployment rate and σ_t^2, we can develop a sense of the changes in wage volatility over the business cycle with the following back of the envelope calculation. Over the course of a typical business cycle in the United States, the unemployment rate fluctuates by roughly 3 percentage points. Our estimate of η_{unemp} then implies that σ_t^2 will increase by 0.013 as the economy moves from the peak of the cycle to the trough.

Many readers will compare our results to the work of Storesletten et al. (2004). Those authors use data on the income of households inclusive of transfers while we use wages of household heads so our results are not directly comparable to theirs. Bearing these differences in mind, one might still ask if our results are plausible in comparison to theirs. Their results imply that the variance of the first-difference in household earnings increases by about 0.041 as the economy moves from expansion to contraction.[7]

To facilitate the comparison of our results to theirs, we also estimate the model using the labor income as data instead of wages. All other features of our analysis remain the same. In this case, the posterior mean of η_{unemp} is 1.14 with a 90% error band extending from 0.76 to 1.54. Performing the same back of the envelope calculation as above produces a difference in income volatility of 0.034 as the unemployment rate increases by 3 percentage points. Given the important differences between their methodology and data and ours, we do not find it surprising that we obtain somewhat different results.[8] In particular, their analysis incorporates information about the Great Depression, which is to say that their sample includes larger fluctuations in the unemployment rate than ours does. In light of this difference, the fact that they find a bigger difference between expansions and contractions is perhaps in line with our finding.

2.4 Impact of Measurement Error

The PSID data on wages are surely affected by measurement error so it is important to consider how our results are affected. Suppose we measure $\mathrm{d}w_{i,t} = \widehat{\mathrm{d}w}_{i,t} + \varepsilon_{i,t}$ where $\widehat{\mathrm{d}w}$ is the true wage growth and ε is measurement error distributed i.i.d. normal with some mean and variance, v_ϵ. Then it follows from equation (1) that $\sigma_t^2 = \hat{\sigma}_t^2 + v_\varepsilon$, where $\hat{\sigma}_t^2$ is the variance of $\widehat{\mathrm{d}w}_{i,t}$. So measurement error of this type will shift the intercept in equation (3), but η_{unemp} is not affected. More generally, if the extent of measurement error varies over time it will affect our result to the extent that it covaries with the unemployment rate.

[7] Appendix C explains how we reach this conclusion from their results.

[8] The difference in methodology is that they look at the cross-sectional dispersion in earnings across cohorts who have lived through different macroeconomic experiences while we look at the dispersion of the *first-differences* in booms and recessions. The difference in data is that they look at total household income inclusive of transfers, while we look at just the labor income of male heads.

2.5 Wage Volatility Over the Business Cycle

One way to assess the role of business cycle fluctuations in driving our finding of a positive correlation between the unemployment rate and wage volatility is to compare years that are close in time but differ in aggregate conditions. To do so, we use the NBER business cycle dates to identify peaks and then choose the years with the lowest and highest unemployment rates between peaks.[9] The pairs of peaks and troughs we identify are (1969, 1971), (1973, 1975), (1979, 1982) and (1989, 1992). Using posterior draws of σ_t^2, we calculate the difference in σ_t^2 for each pair and then compute the average of these differences across pairs. Table 2 shows the distribution of these average differences and the mean and median are both 0.0088. We also explore the increase in wage volatility per unit change in the unemployment rate. For each of the four pairs of years, we compute the ratio of the change in σ_t^2 over the change in the unemployment rate over those years. We then average over the four business cycles. We calculate the posterior distribution of this measure of the slope of the wage-volatility-unemployment relationship. The results in Table 3 show that the mean and median are both 0.31, which is comparable to the estimate of $\eta_{\text{unemp}} = 0.42$ that we obtained from the full sample. Thus we believe that there are important changes in wage volatility over the business cycle.

mean	5%	25%	50%	75%	95%
0.0088	0.0054	0.0074	0.0088	0.0102	0.0122

Table 2: Posterior distribution of difference in σ_t^2 between paired high- and low-unemployment years. Years are selected based on the highest and lowest unemployment rates between NBER peaks.

mean	5%	25%	50%	75%	95%
0.31	0.19	0.26	0.31	0.36	0.43

Table 3: Posterior distribution of $\Delta\sigma_t^2/\Delta u_t$ between paired high- and low-unemployment years. Years are selected based on the highest and lowest unemployment rates between NBER peaks.

[9] As we work with annual data, we treat 1980 to 1982 as a single recession.

3 Decomposition

We now investigate the forces that drive the positive correlation between wage volatility and the unemployment rate. In section 3.1 we lay out our methodology for decomposing the correlation. The first key step in this decomposition is that we use observed covariates to partition the sample into groups within each year. We can then decompose the total variance of wage growth in a year into variances within groups and the variance between groups. Movements in the total variance over years are then driven by movements in these within- and between-group variances as well as shifts in the group sizes. The second key step is to model (statistically) the mean and variance of wage growth within each cell of the partition in each year. Doing so allows us to capture the posterior uncertainty about the components of the decomposition and therefore assess the uncertainty about the decomposition results themselves.

3.1 Methods

Suppose that we can partition individuals into J groups within each year based on observed covariates. When we apply this methodology below we form two groups according to whether an individual has experienced any unemployment spells in the preceding two years and so $J = 2$, but we choose to keep our discussion in general terms to emphasize the fact that this methodology could be applied to any partition of the data. In the sums below, the index j always runs from 1 to J.

3.1.1 Decomposing the Correlation

We now explain our method for decomposing the correlation between wage volatility and the unemployment rate. Given our partition of the data, suppose we know the fraction of individuals in each group and the mean and variance for wage growth in each group. We can then construct the total variance of wage growth across all individuals as

$$\sigma_t^2 = \sum_j \pi_{t,j}\sigma_{t,j}^2 + \sum_j \pi_{t,j}(\alpha_{t,j} - \alpha_t)^2 \equiv W_t + B_t \tag{9}$$

where $\sigma^2_{t,j}$ is the within-group variance and $\alpha_{t,j}$ is the within-group mean for group j at t. Also, $\pi_{t,j}$ is the fraction of observations in group j at time t, and α_t and σ^2_t are the mean and variance for all observations.

As it is customary, we call the two sums in equation (9) *within-* and *between-group* variances. There are two forces that can raise the contribution of the within-group variance. First, the group fractions, $\pi_{t,j}$ might shift towards groups with higher within-group variances. Second, the within-group variances might increase. In order to separate out these two effects, we further decompose the within-group variance term using the cyclical deviations from an overall mean. That is to say, let

$$\bar{\pi}_j = \frac{1}{T}\sum_{t=1}^{T}\pi_{t,j} \qquad \text{and} \qquad d\pi_{t,j} = \pi_{t,j} - \bar{\pi}_j$$

denote the *average* share of observations in each group and the *deviations* from this average, and similarly

$$\bar{\sigma^2}_j = \frac{1}{T}\sum_{t=1}^{T}\sigma^2_{t,j} \qquad \text{and} \qquad d\sigma^2_{t,j} = \sigma^2_{t,j} - \bar{\sigma^2}_j$$

Then we can write W_t as

$$W_t = \sum_j \bar{\pi}_j\bar{\sigma^2}_j + \sum_j \bar{\pi}_j d\sigma^2_{t,j} + \sum_j \bar{\sigma^2}_j d\pi_{t,j} + \underbrace{\sum_j d\pi_{t,j} d\sigma^2_{t,j}}_{e_t} \tag{10}$$

Equation (10) is exact, but we can think of the last term as the second-order error term of a linear approximation, so we will denote it by e_t.

Our goal is to decompose the connection between our cyclical indicator, the unemployment rate, and the overall variance of the wage innovations. We think that this connection is best summarized by the multilevel regression we estimate above, and thus it would be natural to decompose the slope η_{unemp}. Similarly to Gelman (2005), we consider η_{unemp} a *superpopulation* parameter, and work with it's finite population conterpart[10]

$$\hat{\eta}_{\text{unemp}} = \sigma_t^2 \mathbin{/\!/} u_t \tag{11}$$

where $/\!/$ is the univariate regression slope, defined for a time series y_t as

$$y_t \mathbin{/\!/} u_t = \frac{\sum_t (y_t - \bar{y})(u_t - \bar{u})}{\sum_t (u_t - \bar{u})^2} = \frac{\text{Cov}(y, u)}{\text{Var}(u)} \tag{12}$$

Note that $y \mathbin{/\!/} u$ is additive in y,[11] and thus it can be applied to both sides of (9) and (10) to obtain an additive decomposition

$$\sigma_t^2 \mathbin{/\!/} u_t = \underbrace{\sum_j \bar{\pi}_j \left(d\sigma_{t,j}^2 \mathbin{/\!/} u_t \right)}_{\text{cyclical variance of groups}} + \underbrace{\sum_j \bar{\sigma^2}_j \left(d\pi_{t,j} \mathbin{/\!/} u_t \right)}_{\text{composition}} + \underbrace{e_t \mathbin{/\!/} u_t}_{\text{error}} + \underbrace{B_t \mathbin{/\!/} u_t}_{\text{between}} \tag{13}$$

The first term summarizes the effect of the *variance changing within each group* j, this is of course weighted by the proportion of observations within each group. The second term stands for the effect of the number of agents changing within each group: even if we held variances constant, the change in proportions would result in a *composition effect*. We have discussed the error and between group variance terms above, as these turn out to be small they are only included here for the sake of completeness.

We also find it useful to summarize (12) normalized by its left hand side. This provides a very condensed summary of our results that allow the reader to gauge the contribution of each effect, however, when we compare to the moments generated by simulating a calibrated model it turns out to be less useful because our simple model is not able to generate some of the terms.

Our decomposition could be applied directly to sample moments computed from the data. Using sample moments, however, does not give any sense of the uncertainty surrounding the results. To the extent that we partition the sample into small groups, our uncertainty about the sample moments

[10]We find that it makes more sense to compare simulated moments we derive later to finite population their counterparts, also, finite population moments have a lower variance. For a discussion of finite vs superpopulation moments, see Gelman and Hill (2007, Chapter 21.2).

[11]This is easy to check from (12).

can be substantial so it is important to account for this uncertainty. We solve this problem by extending our statistical model to estimate the within-group means and variances. We then draw these means and variances from the posterior distribution and use these draws to compute the terms in equation (13), effectively calculating the distribution of a function of the posterior.

3.1.2 Extending the Model

We now extend our model to the mean and variance of wage growth within each cell of the partition. For an individual i who is in cell j of the partition in year t, the extended model specifies the following distribution for wage growth

$$dw_{i,t} \sim \mathrm{N}\left(X_{i,t}\theta + \alpha_{t,j[i,t]}, \sigma^2_{t,j[i,t]}\right) \tag{14}$$

$$\alpha_{t,j} \sim \mathrm{N}\left(Z_t\kappa_j, \varsigma_{\alpha j}\right) \quad \forall j \tag{15}$$

$$\sigma^2_{t,j} \sim \mathrm{N}\left(Z_t\eta_j, \varsigma_{\sigma^2 j}\right), \quad \sigma^2_{t,j} > 0 \quad \forall j. \tag{16}$$

Equation (14) states that the change in an individual's log wage, $dw_{i,t}$ is normally distributed. As before, the mean of this distribution depends on the individual's demographic characteristics and the coefficients on these demographic characteristics are assumed to be common across groups and across years. In addition, the mean change in wages varies over time and across groups with the $\alpha_{t,j[i,t]}$ term. The notation $j[i,t]$ refers to the group index for the group that individual i is a member of at time t. We also allow the variance of the innovation in wages to vary over time and across groups as captured by the $\sigma^2_{t,j[i,t]}$ term. Equations 15 and 16 model the within group means and variances, respectively. Again, these components are related to aggregate conditions captured by Z_t. Priors are independent for each group, and otherwise have the same form as in section 2.1.1.

3.2 Decomposing by Unemployment Experience

We apply our decomposition method to the PSID data partitioned by unemployment experience. In particular for an individual i at time t, $dw_{i,t}$ refers to wage growth between years $t-1$ and t. We assign an individual to the "no unemployment" group if this individual reports zero hours of unemployment for both years $t-1$ and t. Those reporting positive hours of unemployment in year

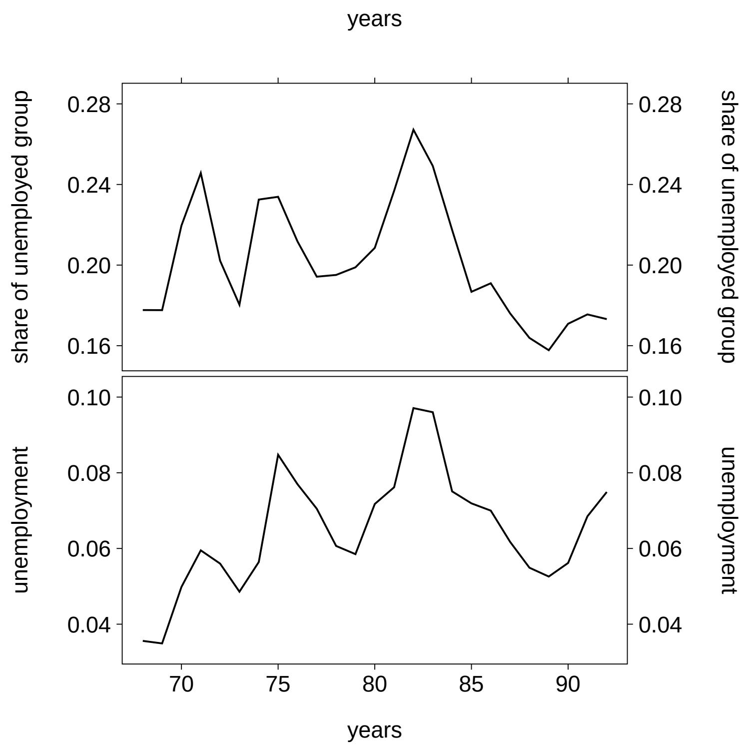

Figure 3: Share of observations in the "unemployment" group (top panel) displayed along with the unemployment rate (bottom panel). As expected, unemployment incidence is highly correlated with the unemployment rate.

$t-1$ or year t are assigned to the "unemployment" group.

As mentioned above, one advantage of the multilevel model is that it is able to partially pool information across years to produce sharper estimates than would be obtained without pooling. The role of equations (15) and (16) is to identify those years for which we expect the α's and σ^2's to be similar. Similarly to section 2, we capture the effect of the business cycle by constructing Z_t from a constant and the unemployment rate.

3.2.1 Decomposition Results

Table 4 shows the posterior distribution of the decomposition (13), in absolute and normalized form.

The first row of the table shows the posterior mean for the fraction of the correlation explained by each component. Here, one can see that countercyclical fluctuations in the within group variances contribute 73% of the correlation with the unemployment group contributing the bulk of this (55%). Most of the remaining correlation comes from the composition effect that arises because the unemployment group has a larger variance at all times and the size of this group increases during a recession. This composition effect contributes 25% of the total. Finally, the error in the decomposition and the between variance contribute next to nothing. The table also shows quantiles for these fractions and one can see that the posterior uncertainty does not change the overall message.

	mean	5%	25%	50%	75%	95%
$d\sigma^2_{\mathrm{U}}$	0.2361	0.1836	0.2147	0.2361	0.2587	0.2900
$d\sigma^2_{\mathrm{E}}$	0.0775	0.0308	0.0574	0.0772	0.0971	0.1244
composition	0.1040	0.0987	0.1018	0.1039	0.1060	0.1095
between	0.0074	0.0014	0.0043	0.0070	0.0100	0.0147
error	0.0018	−0.0075	−0.0019	0.0016	0.0057	0.0113

a) decomposition of regression slopes

	mean	5%	25%	50%	75%	95%
$d\sigma^2_{\mathrm{U}}$	0.5533	0.4732	0.5206	0.5539	0.5866	0.6320
$d\sigma^2_{\mathrm{E}}$	0.1791	0.0810	0.1404	0.1809	0.2193	0.2722
composition	0.2464	0.2060	0.2269	0.2433	0.2626	0.2952
between	0.0174	0.0033	0.0103	0.0165	0.0236	0.0352
error	0.0038	−0.0186	−0.0045	0.0039	0.0131	0.0248

b) normalized decomposition

Table 4: Absolute (top) and normalized (bottom) decomposition of the regression of wage volatility on the unemployment rate. The rows of the both tables are the cyclical variation of the volatility of the unemployed ($d\sigma^2_{\mathrm{U}}$) and employed ($d\sigma^2_{\mathrm{E}}$) groups, followed by the composition effect, and the within and between variance, as shown in equation (13). Note that even though the "unemployed" (U) group makes up only $\approx 20\%$ of the sample, it is responsible for more than half of the cyclicality of the variance. In contrast, the "no unemployment" (E) group contains approximately 4 times that many observations, it only explains abour 1/5 of the cyclicality. Finally, the composition effect explains the rest (about 1/4), and the other terms have a negligible effect.

We view the results in Table 4 as our main empirical results and we now present additional findings from the extended model to explain the forces that drive our results. Figure 4 plots our estimates of $\sigma^2_{t,j}$ for both groups. The top panels show $\sigma^2_{t,j}$ plotted against the unemployment rate and displayed on a common scale. From these plots, one can see that the comovement with the unemployment rate is much stronger in the unemployment group. While the covariance is low in the no unemployment group, this group is about four times the size of the no unemployment group and so this covariance is scaled up by a factor of about 0.8 when the term $d\sigma^2_{t,j}$ is multiplied by $\bar{\pi}_j$ in equation (13) vs 0.2 for the unemployment group.

One way of understanding the benefit of using the multilevel model is to compare the posterior uncertainty of the $\sigma^2_{t,j}$'s to the posterior uncertainty that would result without any pooling (e.g. from a model with a single level — we specify the details of this model in appendix B.3). This comparison shows how much posterior uncertainty is eliminated by partially pooling information across years. Figure 5 displays this comparison, quantifying uncertianty using posterior variances, for each year. For both groups, the multilevel model tightens our estimates of the $\sigma_{t,j}$. The fact that there is a larger benefit of pooling in the no unemployment group reflects the fact that there are fewer observations in this group and so information contained in the hierarchical prior has a larger impact on the posterior.

4 Steady state model and theoretical analysis

In Section 6 we show that a simple, partial equilibrium wage-ladder model with heterogeneous agents and aggregate dynamics can reproduce the empirical facts we document in Sections 2 and 3, and that the most important mechanism behind this result is the movement of the reservation wage in response to job offer and separation probabilities. However, since the calibration of a wage-ladder model requires many parameters (including a distribution for wage offers), it would be difficult to understand the mechanics behind the result and explore discuss the robustness of the calibration relying only on numerical solutions.

However, the following two facts allow us to simplify the analysis of the model considerably:

1. almost all of the movement in the variance of wage innovations comes from changes in the reservation wage.
2. steady state comparative statics give very similar results to the model with aggregate dynamics, and

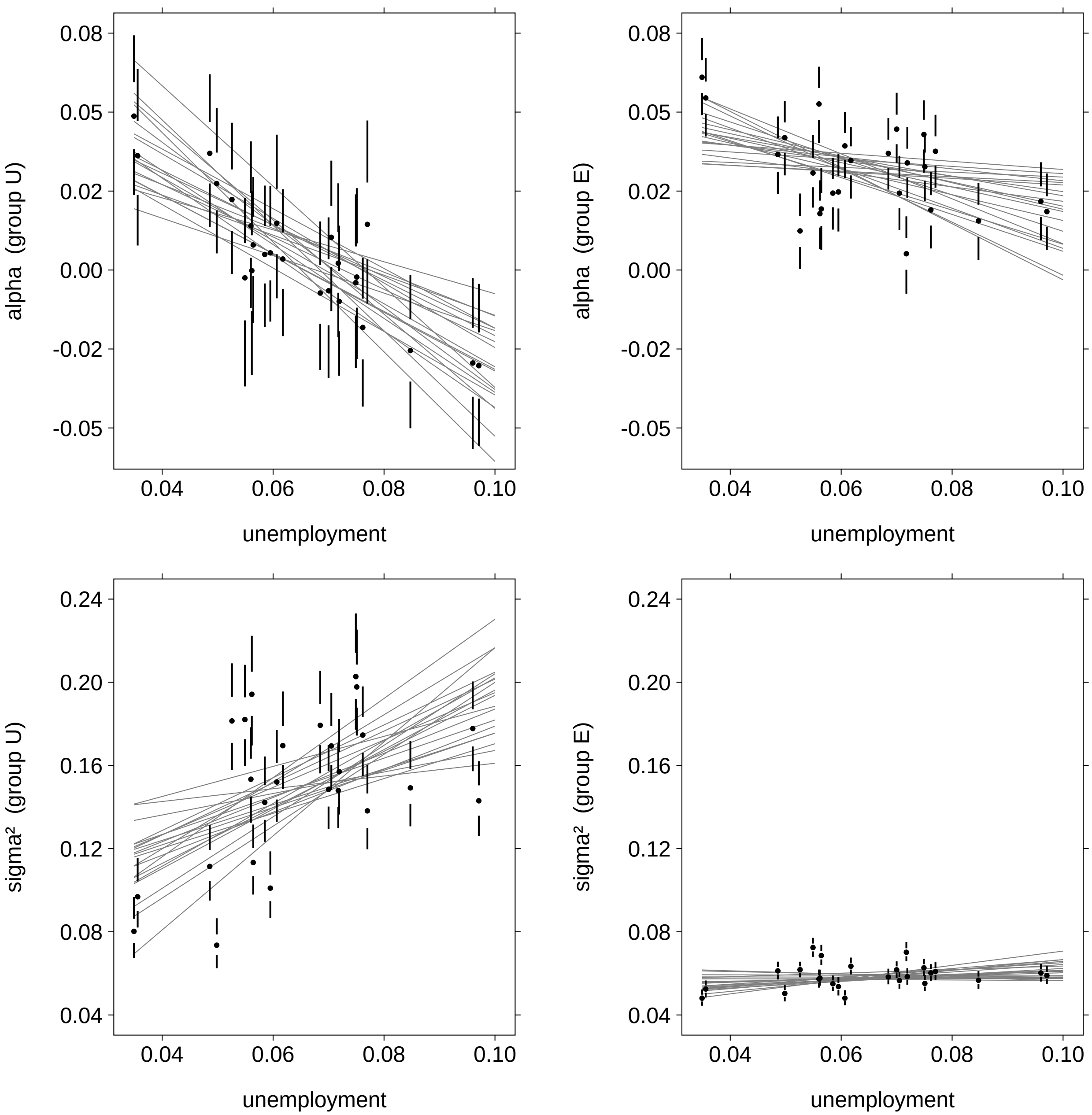

Figure 4: Scatter plot of posterior distribution of α_t (top panels) and σ^2 (bottom panels) versus the unemployment rate in year t, for the "unemployment" (U, left panels) and "no unemployment" (E, right panels) groups. Dots are placed at the posterior medians, vertical bars extend from the 25% to 5% quantiles and from the 75% to the 95% quantiles. The regression lines are random draws from the posterior distribution.

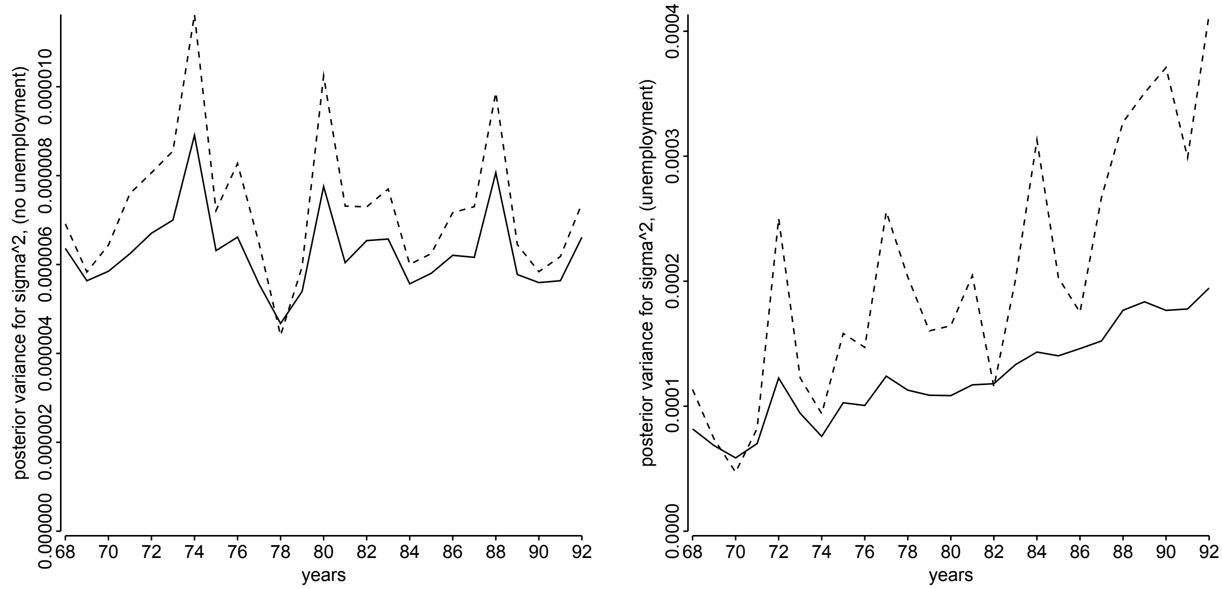

Figure 5: Posterior variance the $\sigma^2_{t,j}$ parameters (by year) for the no unemployment group (left panel) and unemployment group (right panel), in the multilevel (solid lines) and non-pooling models (dashed lines). Clearly, the parameters of interest are estimated with a significantly higher precision in the multilevel model.

3. we can summarize the effect of the job finding and separtes rates in a single parameter.

The first two of these statements are verified in Section 6, in Figures 7 and 10, respectively; we anticipate them here to focus the discussion of the steady state analysis. In this section we set up a model that resembles the one calibrated and simulated in Sections 5 and 6, show that the effects of changes in the job finding and separation probabilities can be summarized by a single parameter (which we denote by χ), obtain analytical approximations for the semi-elasticity of the log reservation wage with respect to χ, and show that this depends on the shape of the distribution via two simple parameters. We also obtain closed-form solutions for the these parameters for the normal distribution.

Consider a partial equilibrium wage ladder model, where the log wage

$$w_{i,t} = p_t + a_{i,t} + x_{i,t}$$

for individual i at time t is the sum of an *aggregate* component p_t, an *individual-specific* component $a_{i,t}$, and a *match-specific* component $x_{i,t}$. We do not specify the firm side of this model, and make the distribution $x_{i,t} \sim F$ exogenous for each firm-worker encounter. Workers have logarithmic utility, and discount factor β. Unemployed workers receive a (log) unemployment benefit $p_t + a_{i,t} + b$. The Bellman equations for the employed and unemployed workers are

$$U(p,a) = p + a + b + \beta\,\mathrm{E}_{(p',a',x')}\Big[\lambda \max\big(W(p',a',x'), U(p',a')\big) + (1-\lambda)U(p',a')\Big] \tag{17}$$

$$\begin{aligned} W(p,a,x) = p + a + x + \beta\,\mathrm{E}_{(p',a',x')}\Big[&\gamma\lambda \max\big[W(p',a',x'), W(p',a',x)\big] + \delta U(p',a') \\ &+ (1-\gamma\lambda-\delta)W(p',a',x)\Big] \end{aligned} \tag{18}$$

where λ and $\gamma\lambda$, are, respectively, the probabilities that an unemployed or employed worker encounters a firm (ie receives a job offer), with γ denoting the relative search efficiency of employed workers compared to the unemployed.[12] Employed workers separate exogenously with probability δ. In Appendix D.1 we show that the value functions separate linearly, and thus we don't specify the stochastic process for p and a at this point. The unemployed workers use a reservation rule, accepting jobs when $x \geq x^*$, where

$$x^* = b + \beta(1-\gamma)\lambda \int_{x^*}^{\infty} \frac{\bar{F}(x)}{1-\beta+\beta\delta+\beta\gamma\lambda\bar{F}(x)} \tag{19}$$

defined x^* and $\bar{F}(x) = 1 - F(x)$ is the tail probability. With some slight abuse of terminology, we refer to x^* as the reservation wage. Introducing

$$r = \frac{1}{\beta} - 1 \qquad \text{and} \qquad \chi = \frac{\lambda}{r+\delta}$$

allows us to rewrite (19) as

$$x^* = b + \int_{x^*}^{\infty} \frac{(1-\gamma)\chi\bar{F}(x)}{1+\gamma\chi\bar{F}(x)}\mathrm{d}x \tag{20}$$

[12] Setting $\gamma = 0$ gives us a variant of the McCall (1970) model.

Total differentiation of (20) and a bit of algebra gives us

$$dx^* = \hat{\mathrm{d}}\chi \frac{\chi}{1+\chi\bar{F}(x^*)} \underbrace{\left(1+\gamma\chi\bar{F}(x^*)\right)(1-\gamma)\int_{x^*}^{\infty}\frac{\bar{F}(x)}{\left(1+\chi\gamma\bar{F}(x)\right)^2}\mathrm{d}x}_{\mathcal{Q}(\gamma)} \tag{21}$$

where we have used $\hat{\mathrm{d}}\chi = d\log\chi$. We derive a first-order approximation for the part denoted $\mathcal{Q}(\gamma)$ by perturbing around $\gamma = 0$:

$$\begin{aligned}
\mathcal{Q}(0) &= \int_{x^*}^{\infty}\bar{F}(x)x^* \\
\mathcal{Q}'(0) &= \left(\chi\bar{F}(x^*)-1\right)\int_{x^*}^{\infty}\bar{F}(x)x^* - 2\chi\int_{x^*}^{\infty}\bar{F}(x)^2x^* \\
\mathcal{Q}(\gamma) &\approx \left(\int_{x^*}^{\infty}\bar{F}(x)x^*\right)\left(1+\gamma\left(\chi\bar{F}(x^*)-1-2\chi\frac{\int_{x^*}^{\infty}\bar{F}(x)^2x^*}{\int_{x^*}^{\infty}\bar{F}(x)x^*}\right)\right)
\end{aligned}$$

We introduce

$$\chi^* = \chi\bar{F}(x^*) = \frac{\lambda\bar{F}(x^*)}{r+\delta} = \frac{\lambda_{\mathrm{UE}}}{r+\delta} \tag{22}$$

which can be calibrated from observed worker flows and the interest rate. Also, let

$$\mathrm{M}_i(F, x^*) = \frac{\int_{x^*}^{\infty}\left(\bar{F}(x)\right)^i\mathrm{d}x}{\bar{F}^i(x^*)} = \int_{x^*}^{\infty}\left(\bar{F}^*(x)\right)^i\mathrm{d}x$$

where

$$\bar{F}^*(x) = \frac{\bar{F}(x)}{\bar{F}(x^*)} \tag{23}$$

which characterizes the tail shape of the distribution: $\bar{F}^*$ is the tail of the *conditional* distribution $x \geq x^*$, and M_i is an expression that depends *only* on the shape of this conditional distribution.

Finally, we arrive at the following approximation for the semi-elasticity of x^* with respect to χ,

$$\frac{dx^*}{\hat{\mathrm{d}}\chi} = \underbrace{\frac{\chi^*}{1+\chi^*}}_{A}\underbrace{\mathrm{M}_1(F,x^*)}_{B}\underbrace{\left(1+\gamma\overbrace{\left(\chi^*\left(1-2\frac{\mathrm{M}_2(F,x^*)}{\mathrm{M}_1(F,x^*)}\right)-1\right)}^{\mathrm{M}_\gamma}\right)}_{C} \tag{24}$$

This is a remarkably simple expression that captures the response of the reservation x^* to χ. From (22)

$$\chi^* = \frac{\lambda_{\mathrm{UE}}}{\delta + r} \approx \frac{\lambda_{\mathrm{UE}}}{\delta} = \frac{1-u}{u}$$

it follows that the first term can be approximated as

$$A = \frac{\chi^*}{1+\chi^*} \approx 1 - u$$

which is easy to calibrate, and it is *independent of the offer distribution* F..

The second term is

$$B = \mathrm{M}_1(F, x^*) = \mathrm{E}_x[x - x^* \mid x \geq x^*]$$

which is the expected gain over the reservation x^* conditional on the unemployed worker accepting the job. This depends on the distribution F, but is *independent of the job finding and separation probabilities.* Thus, if $\gamma = 0$ (no on-the-job search), we can separate the semi-elasticity above into a term that depends on the transition rates, and another that depends on the shape of the distribution of accepted wages. When $\gamma > 0$, the third term C can differ from 1. Since $\bar{F}^* \leq 1$, $M_2/M_1 < 1$, and the slower the tails of the distribution vanish as $x \to \infty$, the smaller this expression is.

We obtain closed-form expressions for $\mathrm{M}_i, i = 1, 2$ for the normal distribution below, but first we show that it is possible to simplify the calibration further by standardizing the distribution of x. Let

$$z = \frac{x - \mu}{s}$$

for some fixed μ and s, and F_z be the distribution transformed from F accordingly. We can transform (20) as

$$z^* = b_z + \int_{z^*}^{\infty} \frac{(1-\gamma)\chi\bar{F}_z(z)}{1+\gamma\chi\bar{F}_z(z)} \mathrm{d}z \tag{25}$$

where

$$b_z = \frac{b - \mu}{s}$$

and we can use the elasticities we have derived, mutatis mutandis. Note that

$$\mathrm{d}x^* = s\,\mathrm{d}z^*$$

and consequently

$$\frac{\mathrm{d}x^*}{\hat{\mathrm{d}}\chi} = s\frac{\mathrm{d}z^*}{\hat{\mathrm{d}}\chi} \tag{26}$$

which allows us to consider the calibration of z^* and s separately.

Now let us assume that $z \sim \mathrm{N}(0,1)$, and introduce

$$\tilde{\mathrm{M}}_i(z^*) = \mathrm{M}_i(\Phi, z^*)$$

where Φ and ϕ are the standard normal distribution and density functions, respectively. Then

$$\frac{\mathrm{d}z^*}{\hat{\mathrm{d}}\chi} = \frac{\chi^*}{1+\chi^*}\tilde{\mathrm{M}}_1(z^*)\left(1+\gamma\overbrace{\left(\chi^*\left(1-2\frac{\tilde{\mathrm{M}}_2(z^*)}{\tilde{\mathrm{M}}_1(z^*)}\right)-1\right)}^{\tilde{\mathrm{M}}_\gamma(z^*)}\right) \tag{27}$$

In Section D.2, we derive the following expressions for the standard normal:

$$\tilde{\mathrm{M}}_1(z^*) = \frac{\phi(z^*)}{\bar{\Phi}(z^*)} - z^* \tag{28}$$

$$\tilde{\mathrm{M}}_2(z^*) = 2\frac{\phi(z^*)}{\bar{\Phi}(z^*)} - \pi^{-1/2}\frac{\bar{\Phi}(\sqrt{2}z^*)}{\bar{\Phi}^2(z^*)} - z^* \tag{29}$$

5 Calibration

Our model has two sources of wage changes: those coming from worker-specific differences captured by the process $a_{i,t}$ and those coming from the search and matching structure, which are reflected in values of $x_{i,t}$. We also allow for some of the variance in wage growth rates that we observe in the data to be driven by measurement error. We assume that $a_{i,t}$ follows an AR(1) process, the parameters of which are stable throughout time; we also assume that the extent of measurement error is stable throughout time. As these components are not time-varying, their main role is to create variation in wage growth rates that is not related to search and matching. We discuss the calibration of these processes in Appendix E.1.

For the search and matching components of our model, our calibration strategy is guided by our theoretical analysis of the steady state reservation wage in Section 4. That analysis shows that the elasticity of the reservation wage with respect to changes in the transition rates depends on (1) the job offer and separation probabilities, via χ^*, (2) the relative search intensity γ of the employed relative to the unemployed, and (3) the shape of the wage offer distribution. Moreover, we have shown when that the latter is standardized, we can separate the effect of the scale parameter and the quantile of the standardized reservation wage $F_z(z)$. We consider the calibration of each of these parameter groups.

The observed job-finding rate is given by $\lambda \bar{F}(x^*)$ and we choose λ to match the job-finding probability reported by Shimer (2007).[13] The separation rate, δ, can be calibrated directly from Shimer's data. We calibrate r to a 4% annual interest rate, these three quantities pin down χ^* using only flow data and the interest rate. We use a result of Nagypál (2005) to calibrate the relative search efficiency γ in a manner that is independent of the distributions, using only the transition probabilities, particularly the job-to-job transition rate as reported by Fallick and Fleischman (2004). The details are discussed in Section E.2.

To pin down z^* (or equivalently, x^*), we choose to calibrate the model so that the fraction of offers that are accepted matches findings from survey evidence. Blau and Robins (1990) use data from the Employment Opportunity Pilot Project baseline survey, which reflected job-search experiences in 1979 and 1980 and found that unemployed individuals received an average of 0.18 offers per week and 10 percent of the unemployed individuals accepted an offer of employment per week. These findings imply that 5/9 of offers received by unemployed individuals are accepted. Krueger and Mueller (2011) conducted a survey of Unemployment Insurance recipients in the state of New Jersey in late 2009 and found that about 60% of job offers are accepted. As the labor

[13]The Shimer (2007) and Fallick and Fleischman (2004) data are monthly transition probabilities and we must convert them to weekly probabilities. Suppose L is a monthly probability, then $\ell = -\log(1-L)$ is the continuous-time arrival rate. $1 - exp(-\ell/4)$ is the weekly transition probability implied by the continuous-time process, which is what we use in our calibration. For the Shimer data, we take the average of this value over the years 1968 through 1992, which corresponds to our PSID sample period.

parameter	description	value
r	interest rate	$(1.04)^{1/48} - 1$
λ	offer arrival rate	0.245
δ	separation rate	0.00978
γ	on the job search efficiency	0.151
ρ	autoregressive coef. for $a_{i,t}$	0.997
σ_a	std. dev. for $a_{i,t}$ innovation	0.000385
s	std. dev. of offer distribution	0.337
$\sigma_{\text{mesasurement error}}$	std. dev. of measurement error	0.148
b	log unemployment benefit factor	-1.619

Table 5: Parameters.

market was particularly slack in 2009, we would expect the steady state acceptance rate to be somewhat lower than 60%. Therefore, we target the $\bar{F}(x^*) = \bar{F}_z(z^*) = 5/9$ as reported by Blau and Robins. Turning to the offer distribution $F(\cdot)$, we assume it is normal and we can normalize the mean of the distribution to zero. We choose the standard deviation s so that the model matches the unconditional variance of wage growth rates in our PSID sample.

In this calibration strategy, we have elected to target $\bar{F}(x^*)$ because our theoretical analysis shows that the results are sensitive to this quantity. As a result, we have the benefit level b pinned down by (19) — note that b does not alter equation (24) except indirectly through x^*. Nevertheless, it is interesting to note that the benefit replacement rate is 13%, quite a bit lower than typical calibration strategies. A low value of b is needed to generate dispersion in wage growth rates: with a dispersed offer distribution, the returns to search are large and in order to prevent unemployed workers from rejecting a large fraction of offers we must make the cost of search large as well, hence a low value of b.[14]

Figure 6 shows the resulting value for the quantities defined in Section 4 and their sensitivity to the calibration of z^*.

[14] This question is discussed in Hornstein et al. (2012).

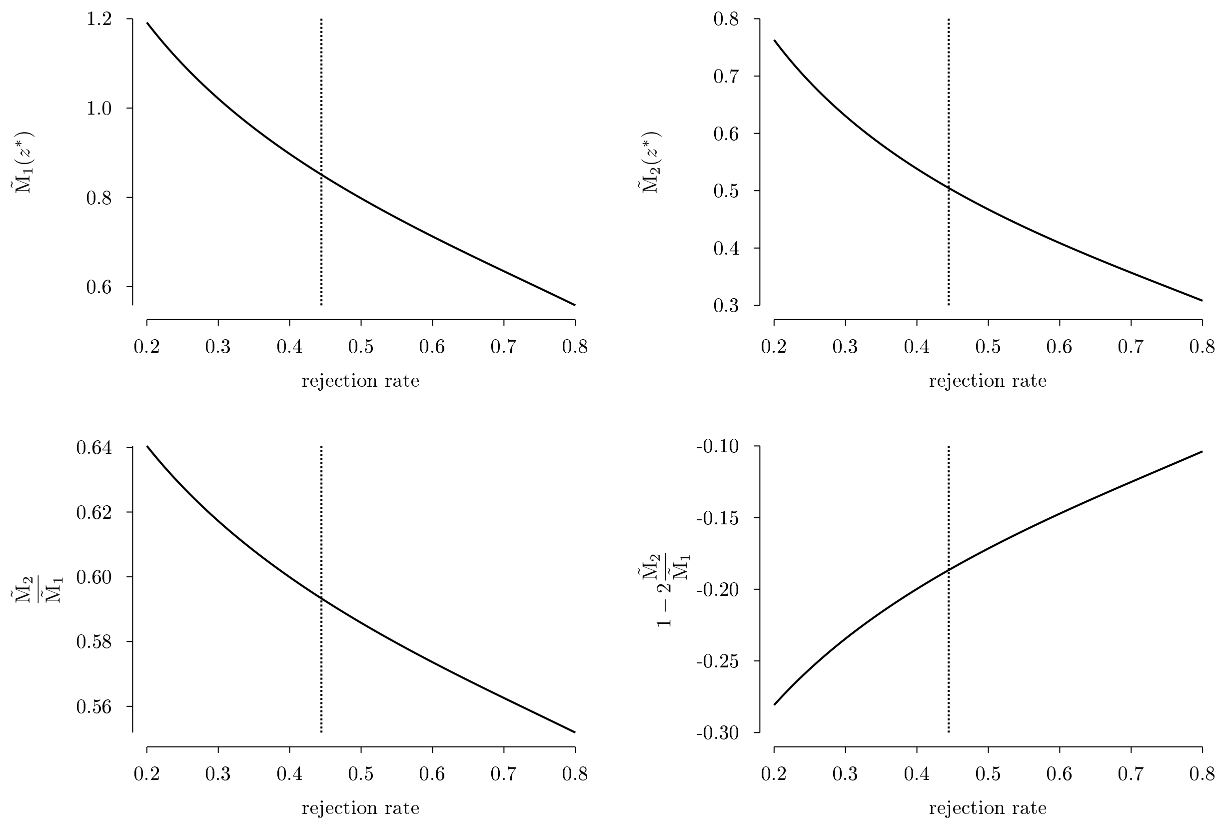

Figure 6: Tail shapes and elasticity components for the normal distribution, as defined and used in equations (27), (28) and (29). The dotted line is at the calibrated z^*, the graphs (most importantly the bottom right one) show that the values are not extremely sensitive to the calibration of the rejection rate.

6 Model with aggregate dynamics

We use our calibrated model to conduct the following experiment: we allow the offer arrival and job-separation rates, λ and δ respectively, to vary over time so that the model reproduces the observed job-finding and separation probabilities from 1948 to 2007 reported by Shimer (2007). When we do this, we assume that all other parameters of the model are constant through time. Changes in the transition rates will, however, affect the worker's choice of reservation wage. In solving the worker's problem, we assume that the worker has perfect foresight for the paths of λ_t

and δ_t.[15] We can then simulate a population of workers from 1948 to 1992.[16] We then compare the wage dynamics from 1968 to 1992 to those we found in the PSID data. Changes in λ_t and δ_t can affect wage dynamics directly (workers move up and fall off the wage ladder more or less quickly) and indirectly through the reservation wage.

In this experiment, we find that the model can generate a substantial fraction of the counter-cyclical volatility in wages, which can be seen in Figure 7. The slope of a linear regression relating the variance of wage growth rates to the unemployment rate is 0.271, while in our empirical estimates of η_{unemp} in Section 2 we found a posterior mean of 0.420. Figure 8 shows a decomposition of this relationship between those experiencing unemployment and those not experiecing unemployment as we did empirically in Section 3. We see here that more than two thirds of the relationship comes from the variance within the unemployment group and the remainder is determined mostly by the composition effect. As we will demonstrate, changes in the variance among those experiencing unemployment are driven by pro-cyclical movements in the reservation wage. The contribution from the variance among those not experiencing unemployment is actually slightly negative as opposed to mildly positive in the empirical estimates. Those who do not go through unemployment are unaffected by changes in the reservation wage and as a result their wage dynamics are only affected by the direct effect of changes in λ_t and δ_t, which we will show is a small effect.

	mean	5%	25%	50%	75%	95%	simulation
$d\sigma^2_{\text{U}}$	0.1913	0.2361	0.1836	0.2147	0.2361	0.2587	0.2900
$d\sigma^2_{\text{E}}$	−0.0212	0.0775	0.0308	0.0574	0.0772	0.0971	0.1244
composition	0.0798	0.1040	0.0987	0.1018	0.1039	0.1060	0.1095
between	0.0227	0.0074	0.0014	0.0043	0.0070	0.0100	0.0147
error	−0.0010	0.0018	−0.0075	−0.0019	0.0016	0.0057	0.0113

Table 6: Estimated and simulated decomposition. The table contains the empirical decomposition from Section 3.2.1 (top panel of Table 4), extended with the simulated decomposition.

As we mentioned, changes in the transition rates have direct effects on the wage distribution and indirect effects through the reservation wage. We perform two alternative simulation experiments

[15] After 2007, we assume that the parameters return to their steady state values. Since we are focussing on the wage dynamics from 1968 to 1992, this return to steady state occurs 15 years after the end of the period we are interested in.

[16] We use the steady state theoretical wage distribution to initialize the model in 1948. The period 1948-1968 acts as burn-in period for our simulation results.

to demonstrate that it is movements in the reservation wage that are driving our results. Let $\{\lambda_t, \delta_t, x^*{}_t\}_{t=0}^T$ be the offer arrival rates, separation rates and reservation wages from our simulation from 1948 to 1992 and let λ, δ, and x^* be their steady state values. First, we simulate the model with the transition rates held constant at their steady state values, but assuming that workers use the reservation wage x^* in period t. The results appear in bottom left panel of Figure 7. The slope of a linear regression fit to these data is 0.319. Second, we conduct the opposite simulation exercise: holding the reservation wage fixed at its steady state value and allowing the transition rates to fluctuate. The results appear in the bottom right panel of Figure 7 and the slope of a linear regression in these data is 0.036. From these simulations, it is clear that movements in the reservation wage are the main source of countercyclical wage volatility in our model. An increase in the reservation wage reduces the dispersion of wage growth rates because when workers transit through unemployment they re-enter employment at a narrower range of wages. Over the cycle, increases in the offer arrival rate and decreases in the separation rate increase the value of searching and raise the reservation wage. Therefore the reservation wage is pro-cyclical, which induces counter-cyclical movements in the dispersion of wage growth rates.

The model generates a plausible level of volatility in the reservation wage. Figure 9 shows the reservation wage relative to its the steady state value. For most of our sample period, the reservation wage is within 15% of the steady state value. Finally, Figure 10 verifies that steady state comparative statics provides a meaningful way to explore the response of the reservation wage.

7 Conclusion

To be written.

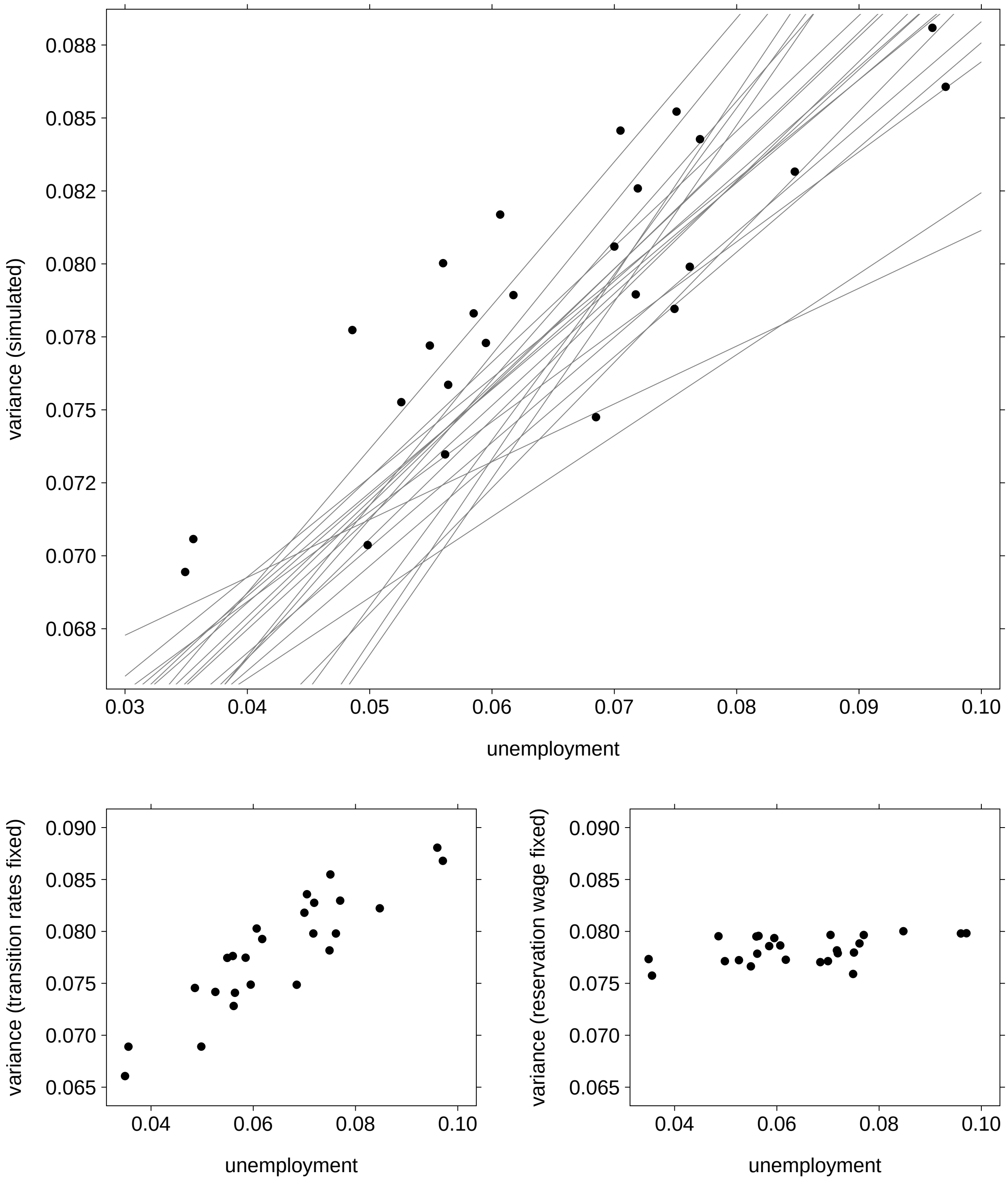

Figure 7: Variance of log wage changes vs unemployment rate for simulated model. Top: variances from simulated model, displayed with posterior draws from posterior of the relevant multilevel regression lines of the empirical estimates in Section 2. Note how the model is able to replicate a significant part of the cyclical variation in the data, but the simulated variance for low values of unemployment is somewhat higher than in the data. Bottom left: simulated series with transition rates fixed at their means, time-varying reservation wage same as the calibrated simulation. Bottom right: reservation wages fixed, transition rates changing. Notice how the the reservation wages generate almost all the variation, not the transition rates.

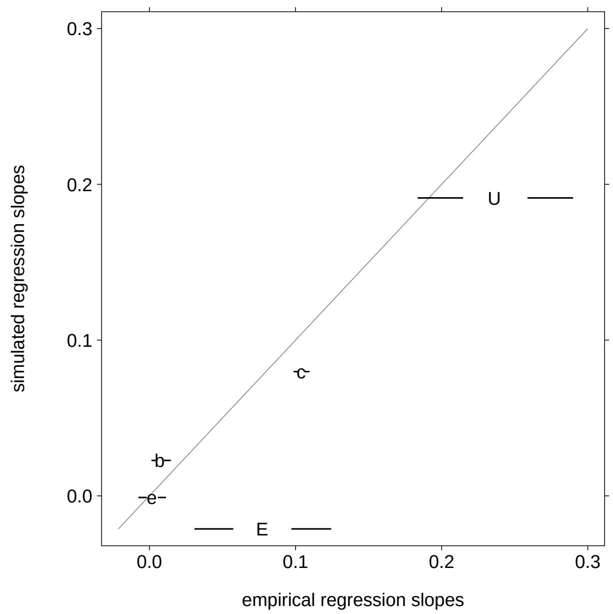

Figure 8: Estimated and simulated decomposition — visualization of Table 6. The horizontal axis shows the empirical quantiles (using the same convention as the rest of the graphs, 5-25-50-75-95% quantiles), while the vertical axis shows the simulated decomposition. The letters in the middle mark the median, with the following legend: (U)nemployed group, no un(E)mployment group, (c)omposition, (b)etween, (e)rror. The grey line is the 45° diagonal. Note that the model comes close to matching all parts of the decomposition, with the exception of group E, which is farther from the diagonal.

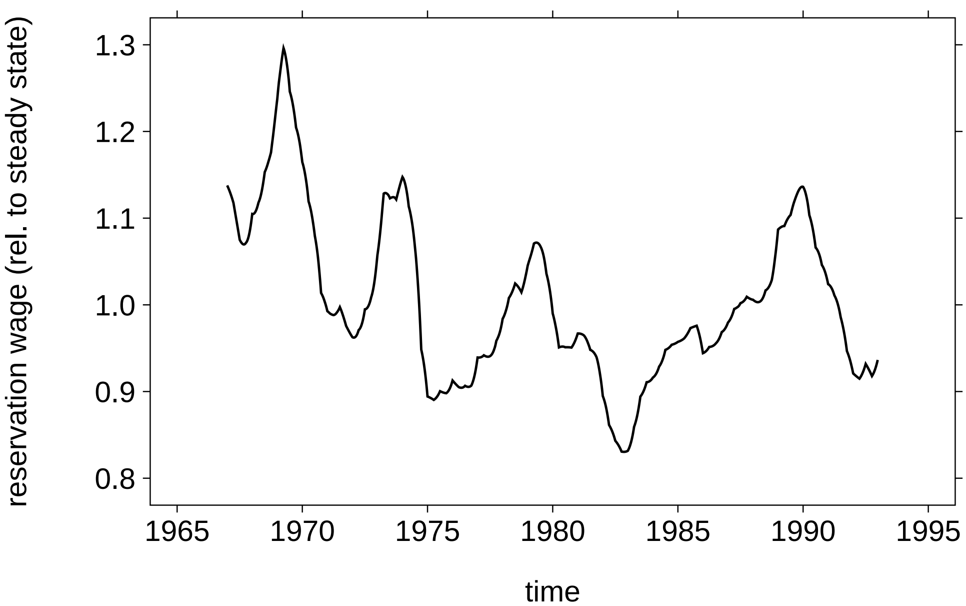

Figure 9: Reservation wage relative to steady state value.

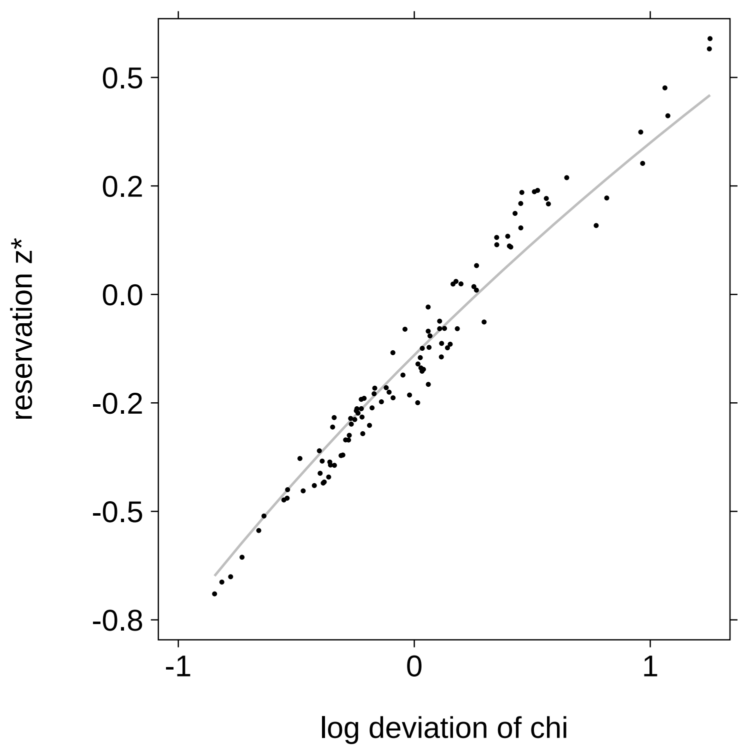

Figure 10: Simulated and analytical reservation wage. The curve shows the locus defined by (25), ie the standardized reservation wage vs the composite index χ in the steady state model, with χ^* and z^* calibrated as in Section 5. The dots show the simulated values of z^* versus the actual χ which are fed into the model. This graph demonstrates that (1) χ is indeed a good summary of the labor market conditions, and (2) steady state comparative statics provides a good approximation to the behavior of the full dynamic model.

Appendix

A Data Appendix

From the PSID, we use data on the annual labor income of male heads and annual hours of work to construct annual wages. In addition, we use data on age and education as covariates. We only include those heads that are between 25 and 60 years of age in both years over which the change in wages is calculated. We drop those with allocated labor income, students, business owners, self-employed individuals, and those with zero hours or income. We trim the top 1% of the income distribution in each year to remove the effect of changes in top-codes across years. Finally, we drop those with wages less than half of the federal minimum wage in that year and we drop those who work fewer than 320 hours in a given year. Our results on income changes in section 2.3 are based on the same sample, but we use annual labor income without dividing by annual hours.

To construct the unemployment-experience partition, we use data on annual hours of unemployment. Let H_t^U be the hours of unemployment in year t. The unemployment group at time t includes those individuals who report $H_t^U > 0$ or $H_{t-1}^U > 0$. The no-unemployment group is those individuals for whom H^U is equal to zero for both years.

For aggregate data, we use the national unemployment rate reported by the BLS and take the average value of the monthly series within each year.

B Methodological Appendix

B.1 The Gibbs sampler

Sampling from the posterior of our model is straightforward and follows Markov Chain Monte Carlo techniques that are commonly used in Bayesian statistics. There are several alternative methods for sampling from the posterior (see Gelman et al. (2004) for a summary). We found that a block Gibbs sampler is relatively fast, easy to implement and has good mixing properties (see Section B.2 for a discussion of convergence). A brief summary of the Gibbs sampler we used is given below.

Our model can be summarized by equations (1)– (3). This model collapses to that in section

2 if one sets $J = 1$ with $j[i,t] = 1$ for all i and t. We partition the parameter space into blocks corresponding to θ, α, κ, ς_α, σ^2, η, and ς_{σ^2}.

B.1.1 Conditional posterior sampling for θ, α, κ, and η

Each of these parameters can be thought of as the coefficient υ in a linear regression

$$c \sim \mathrm{N}(A\upsilon, \Phi) \tag{30}$$

with a known variance matrix $\Phi = \mathrm{diag}(\phi)$ and a given prior $\upsilon \sim \mathrm{N}(\Pi)$. Table 7 shows the mapping between (30) and the model parameters. The conditional posterior distribution is normal, and can be constructed and sampled from in a straightforward manner (Gelman et al., 2004, Sections 14.6 and 14.8).

υ	c	A	ϕ	Π
θ	$\left[\mathrm{d}w_{i,t} - \alpha_{t,j[i,t]}\right]_{i,t}$	$\left[X_{i,t}\right]_{i,t}$	$\left[\sigma^2_{t,j[i,t]}\right]_{i,t}$	flat
α_j	$\left[\mathrm{d}w_{i,t} - X_{i,t}\theta\right]_{i,t:j[i,t]=j}$	$\mathbf{1}$	$\left[\sigma^2_{t,j[i,t]}\right]_{i,t:j[i,t]=j}$	ς_α
κ_j	$\alpha_{\cdot,j}$	Z	$\varsigma_{\alpha j}$	flat
η_j	$\sigma^2_{\cdot,j}$	Z	$\varsigma_{\sigma^2 j}$	flat

Table 7: Conditional posterior sampling for θ, α, κ, and η

B.1.2 Conditional posterior sampling for the group-level variances ς_α and ς_{σ^2}

As is well known, if $h_k \sim \mathrm{N}(0, \varsigma^2)$ (iid, $k = 1, \ldots, K$), then the likelihood is

$$p(\varsigma \mid h) \propto \varsigma^{-K} \exp\left(-\frac{\sum_{k=1}^{K} h_k^2}{2\varsigma^2}\right) \tag{31}$$

As discussed above (see Section 2.1.1), our prior for hyperparameter variances is $p(\varsigma) \propto (\varsigma^2 + s^2)^{-1}$. Then we sample from the conditional posterior using the slice sampler of Neal (2003). Table 8 shows the correspondences between (31) and the model parameters.

h	ς	s
$\alpha_{j,\cdot} - Z\kappa_j$	$\varsigma_{\alpha j}$	s_α
$\sigma^2_{j,\cdot} - Z\eta_j$	$\varsigma_{\sigma^2 j}$	s_{σ^2}

Table 8: Conditional posterior sampling for ς_α, ς_{σ^2}

B.1.3 Conditional posterior sampling for σ^2

Let us fix j and t, and only consider the observations in $I_{t,j} = \{(i',t') : t' = t, j[i',t'] = j\}$. Given (16) and (8), the conditional posterior for $\sigma^2_{j,t}$ is

$$p(\sigma^2_{t,j} \mid \alpha_{t,j}, Z_t, \eta_j, \varsigma_{\sigma^2}) \propto \left(\sigma^2_{t,j}\right)^{-\frac{|I_{j,t}|}{2}} \exp\left(-\frac{\sum_{(i,t)\in I_{t,j}} (\mathrm{d}w_{i,t} - X_{i,t}\theta - \alpha_{t,j})^2}{2\sigma^2_{t,j}}\right) \exp\left(-\frac{(\sigma^2_{t,j} - Z\eta_j)^2}{2\varsigma_{\sigma^2 j}}\right) 1_{\sigma^2_{t,j}\geq 0} \quad (32)$$

This does not correspond to any commonly used family of probability distributions, so we can only sample from it using general tools. After experimenting with rejection methods and obtaining poor acceptance rates, we settled on the slice sampling algorithm of Neal (2003) with excellent results.[17]

B.2 Convergence of the posterior sampler

We monitor the convergence of the Gibbs sampler by calculating the univariate potential scale reduction factor (PSRF) for each parameter value (Gelman and Rubin, 1992; Brooks and Gelman, 1998; Gelman et al., 2004). The PSRF uses variances within and between the parallel chains to estimate the factor by which the scale of the current posterior distribution for a given parameter might be reduced if we were to obtain a sample of infinite size. In practice, values below 1.1 are acceptable, unless very high precision is required. We start 5 chains with overdispersed initial points, calculate the PSRF as the chain evolves for the second half of the chain, and find that it goes below $1.01 - 1.05$ (depending in the parameter) for all parameters after 1000-2000 iterations, which suggests that the mixing is excellent. We generate 5000 parameters for each chain, and discard the first 2500. We found that the mixing was greatly enhanced by subtracting the column

[17] In particular, we used the "stepping out" variant of the algorithm from Neal (2003, Figures 3 and 5).

means from X.

B.3 The non-hierarchical model

We estimate a version of the model without hierarchical regressions for comparison.

$$\begin{aligned}
\mathrm{d}w_{i,t} &\sim \mathrm{N}\big(X_{i,t}\theta + \alpha_{t,j[i,t]}, \sigma^2_{t,j[i,t]}\big) \\
p(\theta) &\propto 1 \\
p(\alpha_{t,j}) &\propto 1 \ \forall t, j \\
p(\sigma^2_{t,j}) &\propto (\sigma^2_{t,j})^{-1} \ \forall t, j
\end{aligned}$$

It is very straightforward to sample from the posterior of this model: first we sample $\theta \mid \alpha, \sigma^2, X$ (see Section B.1.1), then sample from $\alpha_{t,j}, \sigma^2_{t,j} \mid \theta, X$ by sampling from the posterior of the regression of $\mathrm{d}w_{i,t} - \alpha_{t,j[i,t]}$ on 1 with unknown variance and a reference prior for each t, j.

C Comparison to Storesletten et al. (2004)

The purpose of this appendix is to express Storesletten et al.'s results in the same terms as ours to show that the two are not vastly at odds with one another. Throughout, we take their results reported in their Table 2, Panel E because these results are calculated on the assumption that business cycles are defined by the unemployment rate as opposed to GNP growth or NBER cycles and thus closest to our work.

Storesletten et al. specify the following process for the residual of log earnings of individual i

$$\begin{aligned}
u_{it} &= \alpha_i + z_{it} + \varepsilon_{it} \\
z_{it} &= \rho z_{i,t-1} + \eta_{it},
\end{aligned}$$

where $\alpha_i \sim \mathrm{Niid}(0, \sigma^2_\alpha)$, $\varepsilon_{it} \sim \mathrm{Niid}(0, \sigma^2_\varepsilon)$, and $\eta_{it} \sim \mathrm{Niid}(0, \sigma^2_E)$ in an expansion and $\eta_{it} \sim \mathrm{Niid}(0, \sigma^2_C)$ in a contraction. In this notation, our interest is in computing the variance of Δu_{it}, which is

$$\mathrm{Var}\left[\Delta u_{it}\right] = (\rho - 1)^2 \,\mathrm{Var}\left[z_{i,t-1}\right] + \mathrm{Var}\left[\eta_{it}\right] + \mathrm{Var}\left[\Delta\varepsilon_{it}\right]$$

Clearly, as $\rho \to 1$, the first term on the right-hand side goes to zero. This is relevant because Storesletten et al. estimate ρ to be close to 1. For now, suppose $\rho = 1$, but we return to the issue below. Since ε is distributed identically over time, the $\mathrm{Var}\left[\Delta\varepsilon_{it}\right]$ term is a constant. Thus, the difference in $\mathrm{Var}\left[\Delta u_{it}\right]$ between an expansion and a contraction is just the difference in $\mathrm{Var}\left[\eta_{it}\right]$ or $\sigma_C^2 - \sigma_E^2$. Using Storesletten et al.'s estimates, this difference is $0.246^2 - 0.138^2 = 0.041$. In our back of the envelope calculation in section 2.3 we found a difference of 0.013 for wages and 0.034 for earnings.

In the calculation above we ignored the term $(\rho - 1)^2 \,\mathrm{Var}\left[z_{i,t-1}\right]$, which will be countercyclical as the variance of z is countercyclical. As argued above, this term is small. To see this, consider two extreme economies, one that is always in an expansion and one that is always in a contraction. The unconditional variance of z in the expanding economy is then $\sigma_E^2/(1-\rho^2)$ and one can similarly calculate the unconditional variance of z in the contracting economy. The difference between these two variances is an upper bound on the cyclical fluctuations in $\mathrm{Var}(z)$. Using this upper bound, we conclude that the contribution of the term $(\rho - 1)^2 \,\mathrm{Var}\left[z_{i,t-1}\right]$ is at most 0.0013 or 3% of the 0.041 figure we found above.

D Proofs for the steady state analysis

D.1 Linear separability and the reservation wage

Using the independence of the stochastic processes, we rewrite (17) and (18) in terms of

$$W(p, a, x) = W(x) + H(p, a)$$

$$U(p, a) = U + H(p, a)$$

where

$$H(p,a) = p + a + \beta\, \mathrm{E}_{(p',a')}\big[H(p',a')\big]$$

and

$$U = b + \lambda\beta\, \mathrm{E}_{x'} \max\big[W(x'), U\big] \tag{33}$$

$$W(x) = x + \beta\left(\gamma\lambda\, \mathrm{E}_{x'} \max\big[W(x'), W(x)\big] + \delta U + (1-\delta-\gamma\lambda)W(x)\right) \tag{34}$$

Clearly, W is increasing and the unemployed worker follows a reservation rule. Let x^* be the reservation productivity, defined by

$$W(x^*) = U \tag{35}$$

Integration by parts yields[18]

$$\mathrm{E}_{x'} \max\big[W(x'), W(x)\big] - W(x) = \int_{x'}^{\infty} W'(x')\bar{F}(x')x^{*\prime}$$

Then partial differentiation of (34) and the combination of (33) and (34) evaluated at $x = x^*y$ gives us (19).

D.2 Closed-form expressions for the tail characteristics of the standard normal distribution

We continue to use Φ for the CDF of the standard normal. Let $\bar{\Phi}(z) = 1 - \Phi(z)$, and $\phi(z) = \Phi'(z) = -\bar{\Phi}'(z)$. First, we establish the following limit: for all $a \in \mathcal{R}$,

$$\bar{\Phi}(a) = \frac{1}{\sqrt{2\pi}}\int_a^{\infty} e^{-\frac{z^2}{2}}\,dz < \frac{1}{\sqrt{2\pi}}\int_a^{\infty} \frac{z}{a}e^{-\frac{z^2}{2}}\,dz = \frac{1}{\sqrt{2\pi}}\frac{1}{a}e^{-\frac{a^2}{2}}\,dz$$

and thus

$$\lim_{z\to\infty} \bar{\Phi}(z)z = 0 \tag{36}$$

[18] We assume that $\lim_{x\to\infty} \bar{W}(x)F(x) = 0$. Since $W(x)$ is concave, the existence of the first moment of F is sufficient for this to happen.

Then we obtain the following closed-form solutions:[19]

$$\int_{z^*}^{\infty} \bar{\Phi}(z)dz = \underbrace{\left[\bar{\Phi}(z)(z-z^*)\right]_{z^*}^{\infty}}_{=0,\text{ because of (36)}} + \int_{z^*}^{\infty} \phi(z)(z-z^*)dz = \int_{z^*}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} z dz - z^*\bar{\Phi}(z^*)$$

$$= \phi(z^*) - z^*\bar{\Phi}(z^*)$$

$$\tilde{\mathrm{M}}_1(z^*) = \frac{\phi(z^*)}{\bar{\Phi}(z^*)} - z^*$$

and

$$\int_{z^*}^{\infty} \bar{\Phi}^2(z)dz = \underbrace{\left[\bar{\Phi}(z)(z-z^*)\right]_{z^*}^{\infty}}_{=0,\text{ because of (36)}} + \int_{z^*}^{\infty} 2\bar{\Phi}(z)\phi(z)(z-z^*)dz$$

$$= -z^*\bar{\Phi}^2(z^*) + 2\bar{\Phi}(z^*)\phi(z^*) - \int_{z^*}^{\infty} 2\phi^2(z)dz = -z^*\bar{\Phi}^2(z^*) + 2\bar{\Phi}(z^*)\phi(z^*) - \pi^{-1/2}\bar{\Phi}(\sqrt{2}z^*)$$

$$\tilde{\mathrm{M}}_2(z^*) = 2\frac{\phi(z^*)}{\bar{\Phi}(z^*)} - \pi^{-1/2}\frac{\bar{\Phi}(\sqrt{2}z^*)}{\bar{\Phi}^2(z^*)} - z^*$$

E Additional discussion of calibration

E.1 Calibration of the stochastic process for a_t

According to our model, in year t, the log wage of a continuously employed individual is given by

$$w_t = \log\left[\frac{1}{48}\sum_{s=1}^{48} \exp\left(a_t(s) + x_t(s)\right)\right] + \epsilon_t,$$

where s indexes the weeks within year t and ϵ_t is measurement error. If we restrict our attention to those on the same job, then the value of $x_t(s)$ is constant within and across years. As a result, the change in the log wage is given by

$$\Delta w_t = \Delta \log\left[\frac{1}{48}\sum_{s=1}^{48} \exp\left(a_t(s)\right)\right] + \Delta\epsilon_t.$$

[19] The result also follows from Stein's Lemma, eg Lange (2010, p 40).

We will write this as

$$\Delta w_t = \Delta A_t + \Delta \epsilon_t,$$

where A_t is the log-sum term above. A_t will follow a stationary stochastic process that depends on the underlying AR(1) for the a process with parameters ρ and σ.

The autocovariances of Δw_t are

$$\begin{aligned}
\operatorname{Var}(\Delta w_t) &= 2\operatorname{Var}(A_t) - 2\operatorname{Cov}(A_t, A_{t-1}) + 2\operatorname{Var}(\epsilon_t) \\
\operatorname{Cov}(\Delta w_t, \Delta w_{t-1}) &= -\operatorname{Var}(A_t) - \operatorname{Cov}(A_t, A_{t-2}) - \operatorname{Var}(\epsilon_t) \\
\operatorname{Cov}(\Delta w_t, \Delta w_{t-2}) &= 2\operatorname{Cov}(A_t, A_{t-2}) - \operatorname{Cov}(A_t, A_{t-1}) - \operatorname{Cov}(A_t, A_{t-3}).
\end{aligned}$$

We simulate the process $a_t(s)$ and use it to construct A_t. To calibrate the model, we choose ρ, σ, and $\operatorname{Var}(\epsilon_t)$ to minimize the sum of squared differences between the simulated autocovariances and the same moments computed from the data. To compute the moments in the data, we use our same PSID sample,[20] but restrict our attention to individuals with levels of tenure on the job high enough that the differences in wages and the covariances across years involve wages from the same job. For instance, $\operatorname{Cov}(\Delta w_t, \Delta w_{t-2})$ involves $w_t, \cdots, w_{t-3}$ and so we require that the only individuals with job tenure of at least 48 + [interview month] months enters into this calculation. The results are $\operatorname{Var}(\Delta w_t) = .04623$, $\operatorname{Cov}(\Delta w_t, \Delta w_{t-1}) = -.016702$, and $\operatorname{Cov}(\Delta w_t, \Delta w_{t-2}) = -.000987$. We also find the unconditional variance of wage growth rates to be 0.078037.

E.2 Calibration of the relative search efficiency γ

Nagypál (2005) shows that in wage-ladder models it is possible to calibrate the relative search efficiency (or equivalently, the job finding rate) of employed workers merely from the observed transition rates, *without specifying the offer distribution.* We give a concise summary of the method here and show how it applies in our model.

[20] In a first-stage regression we remove age, education and year effects as we do in Sections 2 and 3.

In steady state, the unemployment rate satisfies

$$\lambda \bar{F}(x^*)u = \delta(1-u) \tag{37}$$

where the left and right hand sides are the mass of workers matched or separated each period, respectively. Let $G(x)$ denote the cross-sectional fraction of employed workers with match-specificy productivity below x. Then flow balance requires that

$$\lambda\big(F(x) - F(x^*)\big)u = G(x)\big(\delta + \gamma\lambda\bar{F}(x)\big)(1-u) \tag{38}$$

The left hand side is the flow of workers from unemployment to employment below a particular x — we need to subtract $F(x^*)$ because no offers are accepted below the reservation x^*. The right hand side accounts for exogenous separations and upward transitions on the wage ladder.

We introduce

$$\varphi \equiv \gamma \frac{\lambda \bar{F}(x^*)}{\delta} = \gamma \frac{\lambda_{\mathrm{UE}}}{\delta} \tag{39}$$

which is equal to the ratio of the job finding and separation probabilities, corrected by γ, the relative search efficienty of the employed. All but the last are directly observable from flow data. Below we calibrate φ (and this γ) in a manner independent of the actual distributions F or G, or the reservation x^*. Then we rewrite (38) using (37), (39) and the normalized tail distribution $\bar{F}^*$ (defined in (23)) as

$$1 - \bar{F}^*(x) = G(x)\big(1 + \varphi\bar{F}^*(x)\big)$$

which we can rearrange as

$$G(x) = \frac{1 - \bar{F}^*(x)}{1 + \varphi\bar{F}^*(x)} \tag{40}$$

It follows immediately that the cross-sectional distribution of individual-specific productivities only depends on the transition rates via φ, which we will calibrate direcly from flows. It is useful to rearrange (40) as[21]

$$\varphi \bar{F}^*(x) = \frac{1+\varphi}{1+\varphi G(x)} - 1 \tag{41}$$

We follow Nagypál (2005) in calibrating φ directly. Consider the *average* job-to-job transition probability

$$\lambda_{\mathrm{EE}} = \gamma\lambda \int_{x^*}^{\infty} \bar{F}(x)\mathrm{d}G(x)$$

which we can rewrite as

$$\frac{\lambda_{\mathrm{EE}}}{\delta} = \int_{x^*}^{\infty} \varphi \bar{F}^*(x)\mathrm{d}G(x) = \int_{x^*}^{\infty} \frac{1+\varphi}{1+\varphi G(x)}\mathrm{d}G(x) - 1$$

Using the substitution $y = G(x)$, we arrive at

$$\frac{\lambda_{\mathrm{EE}}}{\delta} = \frac{1+\varphi}{\varphi}\log(1+\varphi) - 1 \tag{42}$$

where the left hand side is observable from flow data, and the right hand side is a monotone increasing function of φ, with limits of 0 and ∞ at $\varphi \to 0$ and $\varphi \to \infty$, respectively.

[21]It is also easy to obtain the inverse from (40): when $G(x) = y$,

$$\bar{F}^*(x) = \frac{1-y}{1+\varphi y} \qquad \text{and thus} \qquad x = \bar{F}^{*-1}\left(\frac{1-y}{1+\varphi y}\right)$$

which is useful for initializing the simulation.

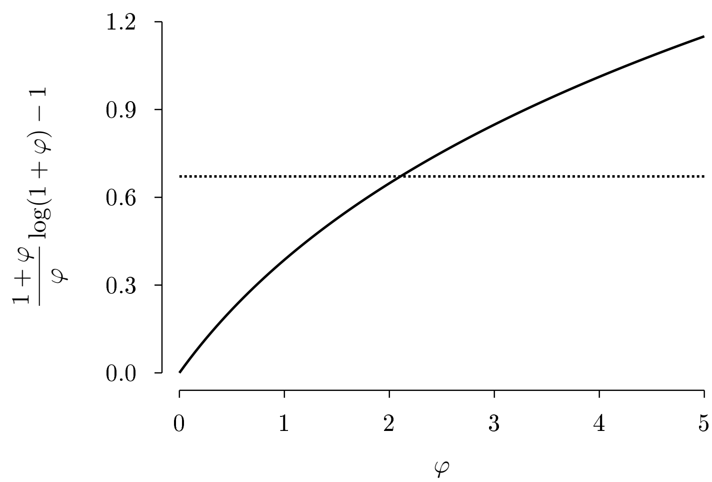

Figure 11: Calibration of φ. The solid line shows the right hand side of (42), while the dotted line shows the target value of $\lambda_{\text{EE}}/\delta$ from Table 5. We solve numerically for $\varphi = 2.106$.

References

Atkeson, A. and Phelan, C. (1994). Reconsidering the costs of business cycles with incomplete markets. In *NBER Macroeconomics Annual 1994, Volume 9*, NBER Chapters, pages 187–218. National Bureau of Economic Research, Inc.

Beaudry, P. and Pages, C. (2001). The cost of business cycles and the stabilization value of unemployment insurance. *European Economic Review*, 45(8):1545–1572.

Bils, M., Chang, Y., and Kim, S.-B. (2011). Worker heterogeneity and endogenous separations in a matching model of unemployment fluctuations. *American Economic Journal: Macroeconomics*, 3(1):128–54.

Blau, D. M. and Robins, P. K. (1990). Job search outcomes for the employed and unemployed. *Journal of Political Economy*, 98(3):pp. 637–655.

Brooks, S. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2):257–73.

Costain, J. S. and Reiter, M. (2005). Stabilization versus insurance: Welfare effects of procyclical taxation under incomplete markets. Technical report, Department of Economics and Business, Universitat Pompeu Fabra.

Costain, J. S. and Reiter, M. (2007). On the relevance of saving for labor market uctuations. Technical report, Bank of Spain and Institute for Advanced Studies.

Fallick, B. and Fleischman, C. A. (2004). Employer-to-employer flows in the u.s. labor market: the complete picture of gross worker flows. Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.).

Gelman, A. (2005). Analysis of variancewhy it is more important than ever. *The Annals of Statistics*, 33(1):1–53.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian analysis*, 1(3):515–533.

Gelman, A., Carlin, B., Stern, H., and Rubin, D. (2004). *Bayesian Data Analysis*. Chapman and Hall, 2nd edition.

Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.

Gomes, J., Greenwood, J., and Rebelo, S. (2001). Equilibrium unemployment. *Journal of Monetary Economics*, 48(1):109–152.

Heathcote, J., Perri, F., and Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the united states: 1967-2006. *Review of Economic Dynamics*, 13(1):15–51.

Hornstein, A., Krusell, P., and Violante, G. (2012). Frictional wage dispersion in search models: a quantitative assessment. *American Economic Review*, page forthcoming. `http://www.econ.nyu.edu/user/violante/Workingpapers/frictional_secrev_aug9_all.pdf`.

Imrohoruglu, A. (1989). Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political Economy*, 97(6):1364–83.

Jung, P. and Kuester, K. (2011). The (un)importance of unemployment fluctuations for the welfare cost of business cycles. *Journal of Economic Dynamics and Control*, 35(10):1744 – 1768.

Krebs, T. (2003). Growth and welfare effects of business cycles in economies with idiosyncratic human capital risk. *Review of Economic Dynamics*, 6(4):846–868.

Krebs, T. (2007). Job displacement risk and the cost of business cycles. *American Economic Review*, 97(3):664–686.

Krueger, A. B. and Mueller, A. (2011). Job search and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data. Technical Report 215, Princeton University Center for Economic Policy Studies.

Krusell, P., Mukoyama, T., and Sahin, A. (2010). Labour-market matching with precautionary savings and aggregate fluctuations. *Review of Economic Studies*, 77(4):1477–1507.

Krusell, P., Mukoyama, T., Sahin, A., and Smith, A. A. (2009). Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12(3):393–402.

Krusell, P. and Smith, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *The Journal of Political Economy*, 106(5):867–896.

Krusell, P. and Smith, A. A. (1999). On the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 2(1):245–272.

Lange, K. (2010). *Applied probability*. Springer Verlag.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *The American Economic Review*, 96(3):461–498.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–26.

Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, 72(1):1–32.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61(3):397–415.

Nagypál, É. (2005). On the extent of job-to-job transitions. `http://faculty.wcas.northwestern.edu/~een461/Extent_job_to_job_09162005.pdf`.

Nakajima, M. (2010). Business cycles in the equilibrium model of labor market search and self-insurance. Working Papers 10-24, Federal Reserve Bank of Philadelphia.

Neal, R. (2003). Slice sampling. *Annals of Statistics*, 31(3):705–741.

Polson, N. and Scott, J. (2011). On the half-cauchy prior for a global scale parameter. *Arxiv preprint arXiv:1104.4937*.

Rudanko, L. (2009). Labor market dynamics under long-term wage contracting. *Journal of Monetary Economics*, 56(2):170–183.

Rudanko, L. (2011). Aggregate and idiosyncratic risk in a frictional labor market. *American Economic Review*, 101(6):2823–43.

Shao, E. and Silos, P. (2007). Uninsurable individual risk and the cyclical behavior of unemployment and vacancies. Working Paper 2007-05, Federal Reserve Bank of Atlanta.

Shimer, R. (2007). Reassessing the ins and outs of unemployment. Technical report, University of Chicago.

Storesletten, K., Telmer, C. I., and Yaron, A. (2001). The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk. *European Economic Review*, 45(7):1311–1339.

Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717.

Western, B. and Bloome, D. (2009). Variance function regressions for studying inequality. *Sociological Methodology*, 39(1):293–326.