# Worker heterogeneity, selection, and unemployment dynamics in a pandemic*

Federico Ravenna† and Carl E. Walsh‡

May 1, 2021

**Abstract**

We employ a new Keynesian model with random search in the labor market and endogenous selection among heterogeneous workers to investigate the impact of a pandemic induced recession on the distribution of unemployment across workers. In such a recession, workers whose unemployment spells in normal times are inefficiently frequent and long are shown to be disproportionately affected. This remains the case even when the pandemic initially causes mass layoffs that affect worker broadly or if many separations take the form of temporary layoffs. Monetary policy that responds to labor market variables affects unemployment for all workers but does little for the distribution of unemployment across the worker types.

**Keywords**: Unemployment, heterogeneity, selection, COVID-19, monetary policy

**JEL classification**: E24, E32, E52.

The COVID-19 pandemic of 2020 generated a severe economic contraction in the global economy, and its impact on unemployment was unlike anything seen in previous recessions. The unemployment rate in the U.S. jumped from 3.5% in February 2020 to 14.8% in April before falling back to 6.7% by December 2020 and 6.0% by March 2021. In contrast, the Great Recession following the global financial crisis resulted in a peak U.S. unemployment rate of 10% in October 2009, which, in turn, was the highest level seen over the previous quarter of a century.[1] It then took until almost 5 years to fall to 6.0%. Using a heterogeneous worker new Keynesian model with search and matching frictions in the labor market, we show how a pandemic recession disproportionately affects those workers who, in normal times, experience longer and more frequent spells of unemployment. Furthermore, we show that endogenous

*We are very grateful for comments from participants at the National Bank of Ukraine Annual Research Conference on "Labor Market and Monetary Policy", 28–29 May 2020, and from Bart Hobijn, Renato Faccini, Yuriy Gorodnichenko, Simon Juul Hviid, Oleksiy Kryvtsov, Karl Walentin, Kenneth West, and two anonymous referees.

†Head of Research, Danmarks Nationalbank, University of Copenhagen, HEC Montreal and CEPR. Email: fera@nationalbanken.dk.

‡Distinguished Professor of Economics Emeritus, University of California, Santa Cruz. Email: walshc@ucsc.edu.

[1] In June 1983, the unemployment rate was 10.1%.

separations and hiring decisions are inefficient in the competitive market equilibrium. Workers with higher average unemployment rates experience more unemployment than is socially efficient; workers with lower average unemployment rates experience less unemployment than is socially efficient. The latter group of workers benefit if separations in a pandemic take the form of temporary layoffs, as layoffs initially did in the COVID-19 recession, but this benefit is not shared by workers with lower average unemployment rates.

To analyze the employment implications of a pandemic recession, we extend a parsimonious model of heterogeneous labor in which firm-worker matches that end are not simply chosen randomly from among existing matches and firms selectively screen job seekers before making hires. We show how selection matters even in a COVID-19 pandemic scenario in which there is a surge in mass layoffs that initially affects all workers non-selectively. Mass layoffs, combined with a fall in aggregate demand, result in a fall in match surplus that results in additional endogenous separations, amplifying the resulting rise in unemployment among those workers whose lifetime labor market outcomes are worse than average. When layoffs are permanent, a COVID-19 pandemic reduces the expected duration of both existing and new matches, lowering the equilibrium return from hiring. These results are reduced but continue to play a role even when the model is parameterized to account for the large rise in the share of temporary layoffs observed in the COVID-19 recession.

The paper makes four primary contributions. First, we identify a new externality when selection arises from labor heterogeneity. Individual firms in the market equilibrium ignore the effects their separation and hiring decisions have on the size and the composition of the pool of unemployed workers. The first effect on the size of the unemployment pool is well-known and gives rise to the Hosios condition for search efficiency. The second effect on the average quality of the unemployment pool distorts the distribution of unemployment across worker types. Second, we model the pandemic as a negative demand shock and a spike in mass layoffs and show how the latter, while initially affecting all workers types, ends up having a disproportional effect on the workers with higher average rates of unemployment. Third, if layoffs in a pandemic are predominately temporary, the recession still induces a rise in endogenous separations, and the employment benefits accrue primarily to the workers with lower average unemployment rates. Fourth, if the central bank responds to labor market variables, it can limit the volatility of unemployment, but the ability of monetary policy to affect the distribution of unemployment across worker types is limited.

Our paper is related to three areas of the literature: research on worker and match heterogeneity, search and matching models with nominal rigidities, and recent work on the macroeconomic effects of COVID-19.

Worker and match heterogeneity play a key role in several models in the search and matching literature and in models with job-to-job transitions (e.g., Guerrieri (2007), Nagypal (2007), Nagypal and Mortensen (2007), Bils et al. (2012), Hall and Schulhofer-Wohl (2018)). Workers differ along many dimensions, and some, such as educational level, specific job skills or

experience, age, and gender may be easily observable. The heterogeneity we focus on arises from *ex ante* unobservable differences among workers.[2] Workers with certain characteristics (young, low-schooling, etc.) experience higher increases in joblessness during a downturn, and several authors (see Grigsby 2020; Baley 2020) find that these workers also differ in unobservable characteristics; ceteris paribus, they have lower productivity. Using CPS data Grigsby (2020) estimates that selectivity in separations and hiring during the Great Recession led to the efficiency of production workers rising by 50%, and to a 10% rise in the aggregate mean human capital of employed workers economy-wide. Positive selection in the employment pool is mirrored into negative selection in the unemployment during recessions. In our model this leads to inefficiency in the allocation and excess volatility of unemployment, relative to an environment not accounting for selection.

In a model with heterogeneous skills and exogenous separation rates, Pries (2008) shows that the composition effect has a large impact on the cyclical value of vacancies and thus on the behavior of employment flows. Ahn and Hamilton (2019) emphasize unobserved differences across workers in (exogenous) unemployment exit probabilities, consistent with the idea that heterogeneity among the workers flowing into unemployment can account for differences in future outflow rates. Kospentaris (2020) argues that unobserved heterogeneity has a large impact on job-finding rates and finds it can explain more than two thirds of total duration dependence in unemployment. This heterogeneity hypothesis (see Davis 1996 and Baker 1992) is central to our approach.

Ravn and Sterk (2017) also develop a model with worker heterogeneity, but they focus on differences in search efficiency rather than productivity differences, and they assume separation probabilities are the same for both types – only job finding rates differ. We allow both separation rates and job finding rates to vary endogenously and to differ across worker types.

While the framework we propose is closely related to this previous work on labor heterogeneity in a search and matching environment, we provide a model with nominal rigidities that allows the role of monetary policy to be analyzed. Our modeling framework is thus part of the literature that combines search and matching labor markets with nominal frictions. Earlier contributors to this area include, among others, Walsh (2003), Trigari (2009), Sala et al. (2008), Thomas (2008), Gertler et al. (2008), and Ravenna and Walsh (2008). These contributions, all assume homogenous workers. We show how layoffs in the competitive equilibrium can be inefficient when labor is heterogeneous, a result that is consistent with that of Berger et al. (2019), who argued for monetary policy to target the layoff rate in a model with countercyclical layoffs.

Finally, a growing number of papers have modelled the macroeconomic implications of

[2] Mincer-wage regressions that condition on observable characteristics of workers exhibit large unexplained residual variation in wages across workers (see Lemieux (2006), and Hornstein et al. (2011)). Other aspects of labor market outcomes are also difficult to explain based on observable worker characteristics. For example, Dickens and Triest (2012) estimate a model of involuntary separation transition probabilities; controlling for age, education, race, and gender, their estimated equation has an $R$-squared of 0.129, suggesting heterogeneity of worker experiences within groups classified based on standard observable characteristics is important.

COVID-19. Guerrieri et al. (2020) focus on sectorial heterogeneity in a two-sector model to show how job destruction in one sector can create a demand-driven recession in the other sector. Kapicke and Rupert (2020) focus on employment adjustments caused by a pandemic within a search and matching framework which distinguishes between workers by health status. Gregory et al. (2020) study the effects of COVID-19 in a search model with worker and sector heterogeneity. They assume transition probabilities between states are partly exogenous, depending on the worker type. Our focus is on how selection affects those probabilities through the impact of shocks on optimal labor market choices. We also incorporate nominal rigidities and endogenous variation in the discount rate, the latter a factor emphasized by Hall (2017) and found by Leduc and Liu (2020) to be important in explaining labor market fluctuations. Consistent with our results and those of Hall (2015) and Ravn and Sterk (2017), the low job-finding rate of some workers plays a crucial role in the behavior of unemployment during recoveries.[3]

The COVID-19 recession has been the result of a variety of underlying shocks that do not easily map into the parsimonious number of shocks typically included in a macro model, and the recent literature has employed different strategies for modelling the causes of the recession. Combinations of supply and demand shocks are employed by Baqaee and Farhi (2020), Fornaro and Wolf (2020), and Kocherlakota (2020). The nature of the supply shock has been treated differently in the literature. Kocherlakota (2020) models it as restrictions on labor supply, while Gregory et al. (2020) assume a temporary decline in worker productivity, and Bernstein et al. (2020) employs a large and persistent job separation shock. Kapicke and Rupert (2020) model an infection shock, and Jackson and Ortego-Marti (2020) combine an infection shock with a skill loss shock that hits the unemployed.

While it is clear that there is not yet a consensus on how to fully replicate the shocks that generated the COVID-19 recession, we choose to combine a negative aggregate demand shock with a shock to exogenous separations that affects all workers as a means of capturing many aspects observed in the pandemic.[4] During the initial stages of the pandemic, stay-at-home measures caused a collapse in demand while lockdowns forced businesses to temporarily close, resulting in a surge in observed unemployment across the entire economy.[5] In contrast to Bernstein et al. (2020) in which all separations are exogenous, we emphasize endogenous separations that amplify the effects of a shock to exogenous separations. We stress how endogenous separations differentially affect those workers that take longer to find new jobs, thus helping to account for the persistence of the rise in unemployment.[6]

[3] In analysing recovery after the Great Recession, Hall (2015) concludes that "The return to normal has been slower than in previous postrecession episodes because the crisis shifted the composition of job seekers toward those with low job-finding rates and low exit rates from unemployment." (p. 121)

[4] Our choice of a separation shock is discussed further in section 3.

[5] Aum et al. (2020) estimate that up to half of job losses in the U.S. and UK may have been due to lockdowns., and the evidence in Kahn et al. (2020) suggests employment losses in April 2020 as measured by unemployment insurance claims were common across U.S. industries and occupations, whether the industry was considered essential or work-from-home capable.

[6] Cheng et al. (2020) finds that "the groups that had the highest unemployment rates in April also tended to

We focus on heterogeneity across workers rather than other dimensions of heterogeneity such as differential effects across sectors or industries. This is motivated in part by recent discussions by monetary policymakers who have displayed interest in the labor market and distributional consequences across workers of monetary policy. For example, Federal Reserve Chair Powell has stressed the gains to those who may benefit from a strong labor market (Powell 2020).[7] Our framework allows us to explore some of the consequences of monetary policy for differences in labor market experiences across workers.

The paper is organized as follows. In section 1, the basic theoretical model of Ravenna and Walsh (2012) on which we build is briefly reviewed. We then show, in section 2, that even when prices are stable and the Hosios condition for search efficiency is satisfied, worker heterogeneity results in inefficient job separations and hiring decisions. In section 3 we use the model to simulate a COVID-19 recession caused by social distancing and lockdown requirements that result in mass layoffs and by a drop in aggregate demand. We then extend the model in section 4 to include temporary layoffs. In section 5 we investigate whether monetary policy responses to labor market variables can reduce inequality in the distribution of unemployment across workers. Conclusions are summarized in section 6.

# 1 The model of productivity heterogeneity

In this section, we describe the basic model of worker heterogeneity and explain how selectivity in hiring and retention decisions affects employment dynamics. The model deviates from a standard NK model with search and matching model with endogenous separations such as Walsh (2005) by adopting the simple model of worker heterogeneity developed in Ravenna and Walsh (2012).[8] As such, we focus here on the key elements that differentiate the model from a basic NK model; details on the complete model can be found in the online appendix. We discuss the baseline selection mechanism assuming only two states for workers - employed or searching for employment; in section 4 we extend the model to allow some unemployed workers to be on temporary layoff and not actively searching for a job.

The model consists of households, wholesale and retail firms, and a monetary policy authority. The representative household purchases consumption goods, holds bonds, and supplies labor to wholesale firms. Wholesale firms hire labor in a market characterized by search and matching frictions and produce a homogeneous good that is sold in a competitive market to retail firms. Retail firms transform the wholesale good into differentiated final goods which

have the lowest reemployment rates, potentially making churn harmful to people and groups with more and/or longer job losses." And Gregory et al. (2020) conclude that "...the lockdown instituted to prevent the spread of the novel coronavirus is shown to have long-lasting negative effects on unemployment. This is so because the lockdown disproportionately disrupts the employment of workers who need years to find stable jobs."

[7] Bergman et al. (2021) use a model of unobserved-heterogeneity across workers to study the effects of monetary policy shock. They show that tight labor markets benefit low-skill workers disproportionally.

[8] Our model is as close as possible to a baseline new Keynesian model of the business cycle with search and matching in the labor market. See Ravenna and Walsh (2012) for more details and for an analysis of the effects of selection on the dynamic impact of productivity shocks.

are sold to households for consumption and to wholesale firms to use in posting job vacancies. Prices of retail goods are sticky ala Calvo.

### 1.1 Worker types and productivity

We assume workers are of two types that differ in their average productivity and its variability: we refer to these types as low-efficiency workers and high-efficiency workers. While an unemployed worker's type is unobserved ex ante, we assume a firm that is hiring engages in a process of interviewing, or screening, during which the firm is able to observe the productivity of a job applicant. Firms can also observe the productivity of their existing employees.[9] Firms employ an (optimal) cutoff productivity strategy; any job applicant whose productivity exceeds the cutoff is hired; any existing worker whose productivity is below the cutoff is fired. This cutoff productivity threshold is endogenous; in a recession it rises so that some unemployed workers who would be hired in normal times are not, and some existing employed workers who would be retained in normal times are not retained.

A fraction $\bar{\gamma}$ of workers are of low ($l$) average efficiency, while the remaining $1-\bar{\gamma}$ are of high ($h$) average efficiency.[10] The worker's efficiency type, $h$ or $l$, is permanently assigned. If $L^j$ denotes the labor force of type $j$, $j = h, l$, we normalize the labor force such that $L^h + L^l = L = 1$. Total employment is $N_t = N_t^l + N_t^h$, where $N^j$ is the number of type $j$ workers who are employed, and $\xi_t \equiv N_t^l/N_t$ measures the fraction of employed workers who are of type $l$. The productivity of a type $h$ worker is constant and equal to $\phi^h$, while the productivity for a type $l$ worker is stochastic and equal to $a_{i,t}^l \phi^l$, where $a_{i,t}^l$ is an idiosyncratic, stochastic productivity shock to worker $i$ of type $l$. We assume $a_{i,t}^l$ is serially uncorrelated and uniformly distributed between zero and one; its cumulative distribution function is denoted by $F(.)$.[11] A low-efficiency worker is less productive than an high-efficiency worker on average, but not always. For high realizations of the idiosyncratic productivity shock, $a_{i,t}^l \phi^l$ can exceed $\phi^h$. The volatility of low-efficiency workers could arise because they could be workers who experience highs and lows, extremely productive at times, unproductive other times, or they may be workers with unstable or chaotic lives outside of work or health issues that get reflected in variation in their job performance. Whatever the source, employers may have difficulty discerning these characteristics of workers without interviewing them or actually observing them as an employee.

[9] By assuming *ex ante unobservable heterogeneity,* the effects we emphasize would still operate within each submarket if labor markets were segmented by observable characteristics.

[10] Ahn and Hamilton (2019) show that the average duration of US unemployment can be matched if the labor force consists of just two types of workers who differ in their job finding probabilities, as will be the case endogenously in our model.

[11] This assumption is for simplicity as it will imply that endogenous separations and interviews that do not lead to hires only involve low skilled workers. In section 2, we discuss the inefficiency of the allocation and the online appendix shows that the efficiency results extend to the case in which both types are treated symmetrically in experiencing idiosyncratic, stochastic fluctuations in productivity and endogenous separations.

## 1.2 Households, labor flows, and vacancies

The household consists of a continuum of workers. The representative household maximizes

$$\mathrm{E}_t \sum_{i=0}^{\infty} \beta^i \left\{ D_t \frac{\mathcal{C}_{t+i}^{1-\sigma}}{1-\sigma} - \left[ v(h_{t+i}^h)(1-\xi_{t+i})N_{t+i} + \xi_{t+i} N_{t+i} \int_{\bar{a}_t}^1 v(h_{i,t+i}^l) f(a) da \right] \right\}, \tag{1}$$

where $\sigma > 0$ is the coefficient of relative risk aversion, $D_t$ is an aggregate preference shock, $\mathcal{C}_t$ is the sum of a market-purchased composite consumption good $C_t$ and home-produced consumption by unemployed workers $C_t^u = (1-N_t)w^u$. In (1),

$$v(h_{t+i}^h)(1-\xi_{t+i})N_{t+i} + \xi_{t+i} N_{t+i} \int_{\bar{a}_t}^1 v(h_{i,t+i}^l) f(a^l) da^l$$

is the disutility to the household of having $N_t$ members working. Hours worked by a type $l$ will depends on the worker's idiosyncratic productivity shock, while because all type $h$ workers are equally productive, they all supply the same number of hours. We assume $v(h_{t+i}) = \ell h_{t+i}^{1+\chi}/(1+\chi)$.

A firm can observe the productivity of its existing employees. However, firms must interview unemployed job applicants to determine a job seeker's current productivity level and efficiency type. The aggregate number of interviews per period is determined through random matching, and all job seekers have identical interview-finding probability, regardless of type. At the interview, the job applicant's productivity level and type is revealed. We assume the (non-stochastic) productivity of an $h$ worker is sufficiently high to guarantee a positive surplus in all states. Thus, if the efficiency type is revealed to be $h$, the worker is hired and produces with probability equal to one. If an interview reveals the job seeker is a type $l$, firms hire only type $l$ workers whose productivity is sufficient to generate a positive surplus. Because currently employed type $l$ workers also receive new idiosyncratic productivity realizations, only those who continue to generate a positive surplus are retained.

At the start of each period, there is an exogenous, stochastic separation probability $\rho_t^x$ that affects all employed workers, regardless of type. We treat the mass layoffs associated with social distancing requirements and lockdowns at the onset of COVID-19 as, in part, an exogenous spike in this separation hazard.

The number of job seekers, denoted by $S_t$, equals those unmatched at the start of the period plus those who do not survive the exogenous separation hazard, or $S_t = 1 - (1-\rho_t^x)\, N_{t-1}$. We define the end-of-period number of unemployed workers as $U_t = 1 - N_t$.[12] Let $S_t^j$ be the number of type $j$ workers who are seeking jobs (so $S_t = S_t^h + S_t^l$) and denote the share of job seekers of type $l$ by $\gamma_t \equiv S_t^l / S_t$. After exogenous separations occur, all other aggregate shocks realizations

[12] The two measures of unemployment can differ as some exogeneously separated workers find employment (and produce) during the period. In search models based on a monthly period of observation, it is more common to assume workers hired in period $t$ do not produce until period $t+1$. Because we base our model on a quarterly frequency, we allow for some workers seeking jobs to find jobs and produce within the same period.

are observed and wholesale firms determine a productivity cutoff $\bar{a}_t^l$ that determines whether a type $l$ worker will generate a positive surplus. The time $t$ idiosyncratic productivity shocks associated with employed low-efficiency workers and with job seekers who are interviewed are observed. With probability $\rho_t^n \equiv F(\bar{a}_t^l)$, a low-efficiency worker's productivity draw will be less than $\bar{a}_t^l$. An unemployed low-efficiency worker with $a_{i,t}^l < \bar{a}_t^l$ who is interviewed is not hired. Absent any direct hiring or firing costs, $\bar{a}_t^l$ is also the cutoff value that determines whether an existing employee is retained.

Three key equations are important for understanding why selection affects unemployment dynamics and the employment experiences of the two types of workers. The first key equation defines the cutoff value of productivity that determines whether a type $l$ work is hired if interviewed and retained if employed. The second key equation relates matching efficiency to vacancies, the number searching workers, and the quality-composition of the pool of unemployed workers. And the third key equation is the job posting condition that links vacancies to the composition of the unemployment pool. We discuss each in turn.

**Cutoff productivity**

Labor is used by wholesale firms to produce a homogenous output that is sold in a competitive market at price $P_t^w$. Let $P_t$ be the final goods price index and define $\mu_t = P_t/P_t^w$ as the retail-price markup. Wholesale firms post vacancies $V_t$, interview and screen applicants, and make retention decisions. The optimization problem of the firm can be written in terms of a key variable, the surplus generated by worker-firm matches, and a critical role is played by the match surplus of a low-efficiency worker. This surplus is equal to

$$s_{i,t}^l = \left(\frac{a_{i,t}^l \phi^l h_{i,t}^l}{\mu_t}\right) - \frac{v(h_{i,t}^l)}{\lambda_t} + q_t^l - w_t^{u,l}, \tag{2}$$

where $h_{i,t}^l$ denotes the hours worked by an employed low-efficiency worker (such a worker produces $a_{i,t}^l \phi^l h_t^l$ of the wholesale good whose real value in terms of retail goods is $a_{i,t}^l \phi^l h_t^l/\mu_t$), $v(h_t^l)$ is the disutility of hours worked, $\lambda_t$ is the marginal utility of consumption, $q_t^l$ is the expected continuation value of a match with a low-efficiency worker, and $w_t^{u,l}$ is the value of an unmatched type $l$ worker's outside opportunity.[13] Hours are chosen to maximize the surplus and thus will vary with a type $l$ worker's idiosyncratic productivity realization. The cutoff value $\bar{a}_t^l$ of a worker's idiosyncratic productivity realization at which $s_{i,t}^l = 0$ is

$$\bar{a}_t^l = \left(\frac{\mu_t}{\phi^l \bar{h}_{i,t}^l}\right)\left(\frac{v(h_{i,t}^l)}{\lambda_t} - q_t^l + w_t^{u,l}\right), \tag{3}$$

where $\bar{h}_t^l$ maximizes the joint surplus for a worker with $a_{i,t}^l = \bar{a}_t^l$.[14] Given household preferences

[13] We assume unmatched workers produce a home consumption good. Details of the surplus derivations are provided in the online appendix.

[14] If we had included an aggregate producitivity shock $z_t$, then the denominator of (3) would become $z_t \phi^l \bar{h}_t^l$ and an increase in $z_t$ would decrease $\bar{a}_t$, increasing hires and retentions.

(1), optimal hours satisfies $v'(h^l_{i,t})/\lambda_t = a^l_{i,t}\phi^h/\mu_t$.[15] Equation (3) implies that $\bar{a}^l_t$ is the same for all firms considering the retention or hire of a low–efficiency worker. An increase in the retail price markup $\mu_t$ reduces the value of intermediate firms' output, and the worker productivity level necessary to generate a positive match surplus rises. This increases $\bar{a}^l_t$, reduces the fraction of low-efficiency job seekers who receive job offers, and increases the endogenous separation rate of already employed low-efficiency workers. This increases the share of low-efficiency workers in the unemployed pool (i.e., $\gamma_t$ rises).

**Efficiency of the matching function**

The number of vacancies posted by wholesale firms $V_t$, together with the number of job seekers $S_t$, determines the number of interviews $I_t$ via a standard CRS matching function:

$$I_t = \psi V_t^{1-a} S_t^a; \qquad 0 < \alpha < 1,\ \psi > 0. \tag{4}$$

A job seeker gets an interview with probability $k^w_t \equiv I_t/S_t = \psi\theta_t^{1-a}$, where $\theta_t \equiv V_t/S_t$. The job finding probability is identical to the interview rate for high-efficiency workers, while for low-efficiency workers it is lower, and equal to $k^{w,l}_t = (1-\rho^n_t)\, k^w_t < k^w_t$. Because the probability a worker drawn from the pool of unemployed job seekers is low-efficiency is $\gamma_t$, the overall job finding probability is

$$k^{w,job}_t = (1-\gamma_t)k^w_t + \gamma_t k^{w,l}_t = (1-\gamma_t\rho^n_t)\, k^w_t. \tag{5}$$

New hires $H_t$ are given by the number of interviewees who are of high-efficiency, all of whom are hired, plus the number of interviewees who are of low-efficiency times the fraction of these with productivity levels that exceed $\bar{a}^l_t$:

$$H_t = (1-\gamma_t)k^w_t S_t + \gamma_t\,(1-\rho^n_t)\, k^w_t S_t = (1-\gamma_t\rho^n_t)\, k^w_t S_t. \tag{6}$$

Screening implies that fewer workers are hired than are interviewed: $H_t < k^w_t S_t$. The number of new hires depends on the endogenous average quality of the pool of unemployed workers as measured by $\gamma_t$ and on the endogenous separation rate $\rho^n_t$. The latter depends on $\bar{a}^l_t$ which, from (3), depends on the retail-price markup. The effective aggregate matching function linking vacancies, job seekers and new hires can be expressed as

$$H_t = (1-\gamma_t\rho^n_t)\, k^w_t S_t = \psi_t V_t^{1-a} S_t^a, \tag{7}$$

where $\psi_t \equiv (1-\gamma_t\rho^n_t)\,\psi < \psi$. Matching efficiency is measured by $\psi_t$. In a recession, both the endogenous separations rate $\rho^n_t$ and the share of type $l$ workers among the pool of job seekers $\gamma_t$ rise, resulting in a fall in $\psi_t$. In a boom, both $\rho^n_t$ and $\gamma_t$ fall. Thus, matching efficiency is endogenous and procyclical.

[15] For type $h$ workers, the condition for optimal hours is $v'(h^h_t)/\lambda_t = \phi^h/\mu_t$, implying $h^h_t$ is the same for all such workers.

**Vacancy posting**

We assume Nash bargaining with firms receiving a share $1-\eta$ of the joint surplus from a match. The job posting condition takes the form

$$k_t^f(1-\eta)\left[(1-\gamma_t)s_t^h+\gamma_t(1-\rho_t^n)\mathrm{E}_t(s_{i,t}^l|hiring)\right]=\kappa, \tag{8}$$

where $\kappa$ is the cost of posting a vacancy, expressed in terms of final goods. The left side of (8) is the probability the firm conducts an interview $k_t^f=\psi\theta_t^{-a}$ times the firm's share of the expected surplus, since with probability $(1-\gamma_t)$ the firm interviews (and hires) a high-efficiency worker and with probability $\gamma_t$, it interviews a low-efficiency worker which results in a hire with probability $1-\rho_t^n$. Because the expected surplus from a high-efficiency worker is greater than the expected surplus obtained from entering into an interview with a low-efficiency worker, the incentive to post vacancies falls when a rise in $\gamma_t$ reduces the average quality of the unemployment pool.

## 1.3 Implications of selection and the transmission of monetary policy

Equations (3) for $\bar{a}_t^l$, (7) for hires, and (8) for job posting are central to understanding the model's key implications that will come into play during a pandemic recession. Consider a negative demand shock that causes a rise in the retail price markup as wholesale prices, which are flexible, fall relative to sticky retail prices. From (3), a rise in the retail price markup $\mu_t$ increases $\bar{a}_t^l$, the critical productivity cutoff for hiring a type $l$ worker or retaining an existing type $l$ employee. This rise in $a_t^l$ increases the endogenous separation rate and generates an inflow into unemployment of type $l$ workers. It also reduces the outflow of type $l$ workers from unemployment as more are screened out in interviews. With the share of low-efficiency workers in the pool of unemployed workers $\gamma_t$ higher, firms posting vacancies are more likely to interview a type $l$ workers and less likely to make a successful hire. From (7), effective hiring efficiency falls and, from (8) the incentive to create vacancies falls. This reduces the job finding probability of type $l$ workers but also of type $h$ workers. Unemployment duration for high-efficiency workers rises, while duration for low-efficiency workers rises both because the probability of getting interviewed has fallen but also because the probability of being hired, conditional on being interviewed, has fallen. These endogenous developments slow the rise of employment during a recession and its subsequent recovery.

A persistent shock to the exogenous separation rate $\rho_t^x$ increases unemployment of both worker types. By reducing the expected duration of matches, the continuation value of a match falls, and, from (3), $\bar{a}_t^l$ rises. This increases endogenous separations, amplifying the rise in unemployment among type $l$ workers. The rise in $\rho_t^n$ due to the rise in $\bar{a}_t^l$, and the rise in $\gamma_t$ as the composition of the job seekers shifts towards low-efficiency workers, reduces the efficiency of the matching process – measured by $\psi_t$ – as selection leads to fewer interviews translating into hires. This also dampens the rise in the job filling rate that occurs as the

number of job seekers rises, which, together with the decline in the expected productivity of job applications as a result of the rise in $\gamma_t$, act to reduce the incentive for firms to post new vacancies.

### 1.4 Retail firms, monetary policy, and market clearing

The rest of the model follows the standard specification in new Keynesian models. The assumption that retail firms adjust their price ala Calvo leads to a basic new Keynesian Phillips curve in which the driver for inflation, real marginal cost, is the price of the wholesale good $P_t^w$, the input of the retail firms, relative to the price of final output $P_t$. Thus, real marginal cost is the inverse of the markup of retail over wholesale goods.

The representative household's first order conditions imply the following intertemporal optimality condition must hold in equilibrium:

$$\lambda_t = \beta(1+i_t)\mathrm{E}_t\left(\frac{P_t}{P_{t+1}}\right)\lambda_{t+1}, \tag{9}$$

where $\lambda_t$ is the marginal utility of consumption and $i_t$ is the nominal rate of interest.

Monetary policy is represented through a simple instrument rule. Our benchmark rule takes the form

$$\ln(1+i_t) = -\ln\beta + \omega_\pi \pi_t. \tag{10}$$

In section 5 we investigate how dynamics are affected if monetary policy also responds to developments in the labor market.

Finally, goods market clearing requires that the household consumption of market-produced goods plus final goods purchased by wholesale firms to cover the costs of posting job vacancies equals the output of the retail sector, or

$$Y_t = \Delta_t\left(C_t + \kappa V_t\right), \tag{11}$$

where $\Delta_t \geq 0$ is a measure of relative price dispersion.[16]

## 2 Inefficient screening and separations

Before carrying out our quantitative exercises of heterogeneity-driven selectivity in hiring and separations in a pandemic-induced recession, we address the implications of worker heterogeneity for the efficiency of the competitive equilibrium. We show that selection leads to a new source of inefficiency; in the competitive equilibrium, individual firms ignore the effects their vacancy posting and separation decisions have on the average quality-composition of the pool of unemployed. Relative to the efficient equilibrium, low-efficiency workers experience spells of unemployment that are inefficiently frequent and average unemployment duration that is

[16] The complete set of equilibrium conditions are given in the online appendix.

inefficiently long. This inefficiency remains even when the Hosios condition is imposed, prices are flexible, and a subsidy to firms offsets the steady-state distortion due to imperfect competition. The efficient equilibrium provides a benchmark for interpreting the distributional consequences across workers of the COVID-19 recession that we analyze in section 3.

In a basic new Keynesian model with search and matching frictions but a homogeneous labor force, Ravenna and Walsh (2011) identify four distortions that arise in the competitive equilibrium with Nash bargaining over wages: (1) a non-zero steady-state markup due to monopolistic competition generates a level of output that is inefficiently low; (2) price rigidity generates inefficient relative price dispersion due to fluctuations in the markup; (3) fluctuations in the markup distort hours from their efficient level; and (4) the vacancy posting condition is inefficient if the Hosios condition is not satisfied. These four distortions would be eliminated if (a) a subsidy to firms is used to raise steady-state output to its efficient level; (b) the markup is constant (i.e., prices are stable); and (c) the Hosios condition holds.

The existence of time-varying worker heterogeneity generates a fifth distortion. When firms separate from low-efficiency employees or screen out such workers at the interview stage, they also jointly determine the average efficiency level of the pool of searching workers from which all new matches are formed. However, firms ignores the impact of their decisions on the size of the unemployment pool and on its quality. The first effect – the externality arising from the impact on the size of the pool of unemployed workers – is eliminated when the Hosios condition is satisfied. The second effect – on the quality of the pool – is not eliminated. The resulting *selection distortion* remains even when prices are stable and the Hosios condition is met.

This distortion generates a wedge between the value of matches in the market equilibrium and in the social planner's problem of maximizing household utility (1) subject to the economy's technology and resource constraints and the search and matching process characterizing the labor market. Let $s_t^h$ and $\bar{s}_t^h$ ($s_{i,t}^l(a_{i,t}^l)$ and $\bar{s}_{i,t}^l(a_{i,t}^l)$) denote the joint surplus for a match with a high-efficiency (low-efficiency) worker in the market equilibrium and the social planner's allocation, respectively. For matches with a low-efficiency worker, the surpluses will depend on the workers idiosyncratic productivity level $a_{i,t}^l$. Define $\mathcal{S}_t^h \equiv \bar{s}_t^h - s_t^h$ and $\mathcal{S}_{i,t}^l(a_{i,t}^l) \equiv \bar{s}_{i,t}^l(a_{i,t}^l) - s_{i,t}^l(a_{i,t}^l)$. Evaluated at the efficient equilibrium, the online appendix shows that

$$\mathcal{S}_t^h = \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right) (1-\rho_{t+1}^x)\left(1-\alpha k_{t+1}^w\right) \mathcal{S}_{t+1}^h - \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right) \gamma_{t+1} X_{t+1} \tag{12}$$

and

$$\begin{aligned} \mathcal{S}_{i,t}^l(a_{i,t}^l) \quad = \quad & \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right) (1-\rho_{t+1}^x)\left(1-\alpha k_{t+1}^w\right)\left(1-\rho_{t+1}^n\right) \mathcal{S}_{i,t+1}^l(a_{i,t+1}^l) \\ & + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right) \left(1-\gamma_{t+1}\right) X_{t+1}, \end{aligned} \tag{13}$$

where

$$X_{t+1} \equiv (1-\alpha)\left(1-\rho^x_{t+1}\right) k^w_{t+1}\left[\bar{s}^h_{t+1} - \left(1-\rho^n_{t+1}\right)\bar{s}^l_{i,t+1}(a^l_{i,t+1})\right]. \tag{14}$$

The sign of $X_{t+1}$ depends on the expected value of $\bar{s}^h_{t+1} - \left(1-\rho^n_{t+1}\right)\bar{s}^l_{t+1}(a^l_{i,t+1})$, the average surplus at $t+1$ of a high-efficiency worker, conditional on surviving the exogenous separation hazard, minus the average surplus at $t+1$ of a low-efficiency worker, conditional on that worker surviving the exogenous separation hazard and being retained.[17]

The market equilibrium is efficient if and only if $\mathcal{S}^h_t = \mathcal{S}^l_{i,t}(a^l_{i,t}) = 0$ for all $i$ and $t$. For example, if labor is homogeneous and all workers are high-efficiency, only (12) is relevant, $\gamma_{t+1} = 0$ (there are no low-efficiency workers among the unemployed), and $\gamma_{t+1}X_{t+1} = 0$. In this case, (12) becomes $\mathcal{S}^h_t = \beta \mathrm{E}_t \Lambda_{t+1}\mathcal{S}^h_{t+1}$ which is satisfied if $\mathcal{S}^h_t = 0$ for all $t$. Similarly, if all workers are low-efficiency, $\gamma_{t+1} = 1$ and $\mathcal{S}^l_{i,t}(a^l_{i,t}) = 0$.

In the presence of labor types of differing efficiency, both $\mathcal{S}^h_t$ and $\mathcal{S}^l_{i,t}(a^l_{i,t})$ will differ from zero. The terms involving $X_{t+1}$ appear because the social planner accounts for the effect of worker type on the composition of the unemployment pool, an effect ignored by firms in the competitive equilibrium. Consider first the case of a type $h$ worker. The social planner internalizes the effect the employment of an additional high-efficiency worker has in lowering the average productivity of the pool of job seekers, making it more likely that a new hire would be a low-efficiency worker. This "cost" is measured by the last term in (12) and implies $\mathcal{S}^h_t = \bar{s}^h_t - s^h_t < 0$; it reduces the surplus of a type $h$ worker from the perspective of the social planner relative to the firm's valuation.

Hiring a type $l$ worker improves the average productivity of the remaining pool of unemployed workers, and, from the perspective of the social planner, this increases the valuation of matching with a type $l$ worker. The social planner places a higher valuation on a match with such a worker relative to the firm's valuation so $\mathcal{S}^l_{i,t}(a^l_{i,t}) \equiv \bar{s}^l_{i,t}(a^l_{i,t}) - s^l_{i,t}(a^l_{i,t}) > 0$.

Thus, in the market equilibrium, firms over value high-efficiency workers and under value low-efficiency workers relative to the social planner. The lower valuation placed on a match with a low-efficiency worker implies that the cutoff productivity level for hiring and retaining workers in the market equilibrium is too high. Low-efficiency workers who experience endogenous separation and become unemployed in the competitive equilibrium would remain employed by the social planner. Similarly, some low-efficiency workers who obtain interviews but are screened out in the competitive equilibrium would be hired by the social planner. As a result, low-efficiency workers face unemployment spells that are too frequent and too long.

This also translates into a higher share of low-efficiency workers among the unemployed and a lower expected benefit to posting vacancies in the market equilibrium. Reduced job posting implies high-efficiency workers also experience a lower job finding rate and longer average duration of unemployment. Ceteris paribus, endogenous separations are too high in the competitive equilibrium, average unemployment is also too high, and average unemployment

[17] $X_{t+1}$ does not depends on $i$ because $s^l_{i,t}$ is i.i.d. $X_{t+1}$ is a function of the two labor types but not on the idiosyncratic realizations of $a^l_{i,t+1}$.

duration is inefficiently long.[18]

In the next section, we use the efficient allocation as a benchmark for studying the market response to a COVID-19 recession.

# 3 A COVID-19 recession and the role of worker heterogeneity

We use the model of worker heterogeneity to investigate the impact of a COVID-19 recession and discuss the key propagation mechanism that cause the pandemic to have differential effects across worker-types. In this section we assume that firms can only recruit from the unemployment pool and workers can only be employed or searching for work. Our baseline model captures the recession and recovery path in an economy with permanent separations and where labor separations require that firm-worker matches be re-established through a costly matching process. In section 4 we allow for some workers to be on temporary layoff and therefore unemployed but not actively searching for work.

To begin, we discuss the parameterization and the shocks we employ to capture the COVID-19 recession. We then explore the responses implied by the market equilibrium and compare these to the efficient responses. Our main results are: (1) low-efficiency workers are disproportionately affected in a pandemic recession even though the initial spike in job loss affects both worker types; and (2) responses in the market equilibrium are inefficient, and the labor market experiences of those workers who generally experience poorer labor market outcomes worsen significantly more, relative to the efficient equilibrium, as compared to high efficiency workers.

## 3.1 Parameterization

Six parameters are key to determining the impact of labor-productivity heterogeneity on the model economy. These parameter are the value of home production $w^u$, the coefficient $\ell$ scaling the disutility of labor hours, the cost of vacancy posting $\kappa$, the productivity of the matching technology $\psi$, the relative steady-state productivity of high- to low-efficiency workers $\phi^h / \left(\phi^l \int_0^1 a_i^l dF(a_i^l)\right)$ and the labor force share of low-efficiency workers $\bar{\gamma}$. Values for these parameters are selected by jointly targeting six steady-state values: the steady-state aggregate unemployment rate $U_{ss}$ and the unemployment rates $U_{ss}^l$ and $U_{ss}^h$ for each worker type, average hours per worker $h_{ss}^{av}$, vacancy posting costs $\kappa V_{ss}$ as a share of output, and the probability $k_{ss}^f$ of a vacancy match with a job application. The target steady state values and the implied parameters are reported in Table 1.

$U_{ss}$ is set to the average U.S. civilian unemployment rate over 1948:Q4 to 2019:Q4. Neither $U_{ss}^l$ nor $U_{ss}^h$ are directly observable, so our baseline parameterization follows Gregory et al. (2021) who use the Longitudinal Employer and Household Dynamics (LEHD) data from 1997

[18] The appendix shows that this result can be extended to the case in which both worker types experience individual-specific i.i.d. productivity shocks and endogenous separations.

to 2014 to estimate the labor market shares and unemployment rates for three separate worker types that differ by employment duration. We map these estimates into our model with two types of workers, implying a share of low-efficiency workers in the labor force of 38% with an average unemployment rate of 9.87%, and a share of high-efficiency workers of 62%, with an average unemployment rate of 2.97%.

This baseline parameterization implies an unemployment rate ratio $U^l_{ss}/U^h_{ss}$ equal to 3.3. Given that $U^h_{ss}$ and $U^l_{ss}$ are not observable, for robustness we considered alternatives that resulted in a higher value of 4.2 and a lower value of 2.5 for this ratio, keeping $U_{ss}$ constant at 5.6%. The higher value of $U^l_{ss}/U^h_{ss}$ comes from an alternative mapping of the three labor types in Gregory et al. (2021) into our model with two labor types that implies $\bar{\gamma} = 20\%$, $U^h_{ss} = 2.1\%$ and $U^l_{ss} = 14.4\%$. The lower value of $U^l_{ss}/U^h_{ss}$ is based on the observed difference in average unemployment rate for workers aged 16 to 24 (4.4%) and over-24 (11.6%). Our general conclusions are robust to these alternative parameterizations and results are reported in the online appendix.

The other targeted moments reported in Table 1 include steady-state hours per worker $h^{av}_{ss}$, the steady-state aggregate separation rate $\rho_{ss}$ and the probability of a match between an applicant and a vacancy $k^f_{ss}$. These values are parameterized to standard values in U.S. business cycle literature. The share of output devoted to hiring activities is in line with empirical evidence reported in Ravenna and Walsh (2008). For parameters standard to new Keynesian models, we adopt values common in the literature, and for the benchmark monetary policy rule, we use a standard value of 1.5 for the response coefficient on inflation.

Table 1 summarizes the key parameter and steady-state values. In our parameterization, the share of type $l$ workers in the total labor force is 38%. Because the separation rate for these workers is about 40% larger than the average separation rate, their share in the steady-state pool of job seekers $\gamma_{ss}$ is 52%, while their share $\xi_{ss}$ in steady-state employment is only 36%.[19] Thus, low-efficiency workers are over-represented in the pool of unemployed, and this pool has a lower average productivity than the pool of employed workers, as reported in table 1. When matched for an interview with a firm, high-efficiency workers are expected to have an hourly productivity 16% higher than low-efficiency workers. We assume the i.i.d. productivity shock $a^l_{i,t}$ has a uniform distribution with support $(0, 1]$.The productivity ratio between employed and unemployed workers is smaller, and equal to 1.04, since only relatively highly productivity type $l$ workers are retained in employment. The extent of selection at hiring is small; firms screen out only about 2.55% of the workers they interview and just 4.9% of the type $l$ workers who are interviewed. However, as (7) showed, the screening-out rate increases in a recession as $\gamma_t$, the share of low-efficiency workers among the unemployed increases and $\bar{a}^l_t$, the minimum productivity for a match to generate a positive value, increases.

In the simulations, we follow the standard approach in NK models of assuming the existence

[19] The value for excess-separation rate of type $l$ workers $\rho^{ss}_n$ is the consequence of a parameterization requiring both a high ratio between unemployment between type $l$ and $h$ workers and a high quarterly probability of an interview $k^f_{ss} = 0.9$.

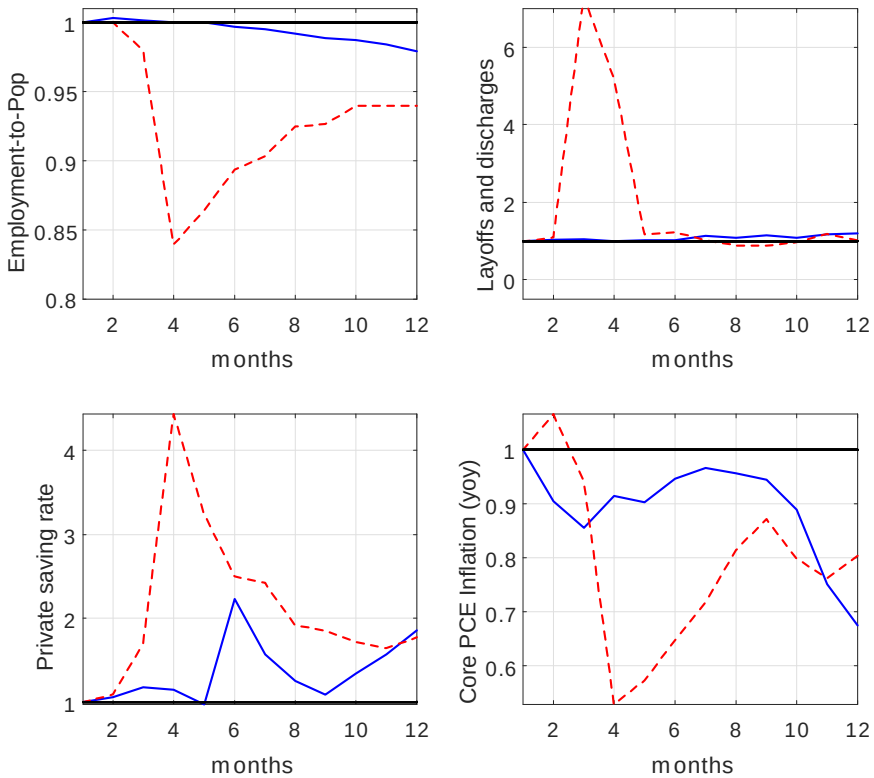


Figure 1: Macro developments during 2020 in the U.S. The time scale denotes months since December 2007 (solid lines) for the Great Recession and January 2020 (dashed lines) for the COVID recession.

of a tax/subsidy system that ensures the steady-state allocations in the planner's problem and the market allocation coincide. This requires that the Hosios condition holds and that the study-state distortions due to imperfect competition and the sectional distortion are eliminated.

### 3.2 Shocks

Figure 1 provides some macroeconomic evidence on the distinctive character of the COVID-19 recession. It illustrates the behavior of the employment-to-population ratio, layoffs and discharges, the private saving rate, and inflation over the first 12 months of the COVID-19 recession and, for comparison, the first 12 months of the Great Recession following the global financial crisis. All values are scaled to unity at the start of the recession. The spike in layoffs and discharges measured by JOLTS data during COVID-19 contrasts sharply with their behavior in the Great Recession, during which the initial rise in layoffs is barely discernible and where layoffs then rose gradually over the first year of the recession. In the COVID-19 recession, shelter-in-place orders, social distancing requirements and breakdowns in supply chains reduced the economy's ability to produce without generating a corresponding drop in demand and, as such, were associated with a small rise in inflation.

While sectorial redistributional effects were important, the rise in the saving rate and the subsequent fall in inflation shown in the figure suggest the pandemic also featured an overall

drop in spending, as would be generated by a negative aggregate demand shock.

To describe the COVID-19 recession, therefore, we assume that the economy is hit by both a positive shock to exogenous separations and a negative demand shock. Rather than simulate the model's response to exogenous shocks as in a standard impulse response exercise, we build a hypothetical scenario based on observable variables, specifically total separations and output, and then let the model back out the demand and separation shocks that drive the dynamics. The model uses the data to endogenously allocate the path of total separations across an exogenous, unexpected shock to $\rho_t^x$, an endogenous portion driven by selection, and an exogenous demand shock. Similarly, the model will back out the exogenous shift in household preferences and separations consistent with the path for output. We then produce conditional forecasts for labor market variables and other aggregate variables which are not supplied to the model.

Based on CPS data, Cortes and Forsythe (2020b) report that over 80% of the pandemic-induced fall in employment in April 2020 resulted from individuals exiting employment, while reduced hiring accounted for the rest. The public heath responses to COVID-19, including business closures, lockdowns and social distancing requirements, resulted in the destruction of job matches and had features common to a negative supply shock in that it reduced the economy's ability to produce without generating a corresponding drop in demand. The separation shock $\rho_t^x$ captures some of the supply side aspects of COVID-19 and adversely affects all workers, regardless of type. Guerrieri et al. (2020) model COVID-19 as a shock that destroys jobs, much as our separation shock does. Gregory et al. (2020) models the decline in employment as a negative aggregate TFP shock. Such a shock would be expected to lead to an increase in separations, a fall in hiring, and a decline in wages for workers who remain employed. However, Cortes and Forsythe (2020a) find that the decline in aggregate employment, rather than a decline in average wages accounts, for the observed decline in labor earnings. We then capture the accompanying fall in aggregate demand through a negative preference shock $D_t$, representing a shift in households' utility away from current consumption and towards home-production. Both shocks will generate further responses in total separations as endogenous separations adjust.

To produce the COVID-19 scenario, we condition on forecast paths for output and total separation produced at the beginning of the COVID-19 recession. This means our simulations are hypothetical paths in the absence of the effects of the vast income support measures (the CARES act) and expansionary monetary policies (discretionary cuts in the federal funds rate, forward guidance, and quantitative easing measures) which are outside our model and that affected the actual path of the economy within the first months of the pandemic. Our target output path is based on the 2020Q2 vintage of the Survey of Professional Forecasters (SPF) for GDP growth, relative to a potential growth path of 2% per year. The SPF predicts GDP growth only up to 5 quarters ahead. For the subsequent 5 quarters, we assume the economy recovers at the same rate as over the forecast horizon, with the shortfall in output dissipating at

an estimated quarter-over-quarter rate of 67.4%. The SPF does not provide a forecast for total separations, so our target path for separations is obtained by measuring the 2020Q2 increase in the cumulated CPS transition rate from employment to non-employment over each month, as a share of the previous month employment stock, relative to the transition rate averaged over the 2018-2019 period.[20] The path for total separations then assumes the initial rise reverts back to steady state at the same rate as the SPF output forecast.[21]

Both shocks are essential to discuss a pandemic-induced recession. The exogenous separation shocks act as a supply shock; the value of each existing or potential match falls, ceteris paribus, since the cost per vacancy is fixed but its return in terms of match-lifetime production is lower, given that the match is expected to have a shorter duration. By itself, the separation shock is not sufficient to lower demand in line with the SPF forecast for the U.S. economy as it pushes firms to quickly rehire a large share of the separating workers. The preference shock acts as a demand shock that allows the model to match the output loss for a given separation path. The demand shock lowers output and the demand for labor but cannot, by itself, reasonably account for all of the surge in total separations. Together, the two shocks capture important aspects of the 2020 recession.

### 3.3 Mass layoffs and the distribution of unemployment across worker types

Our first set of results explains why selection changes the dynamics of the economy in the COVID-19 recession and reallocates the burden of the recession across worker types. On the workers' side, we find that low-efficiency workers are disproportionately affected in a pandemic recession even though the initial spike in job loss affects both worker types. We also find that selection worsens labor market conditions for *all* worker types. On the firms' side, we show that the distance between the vacancy-filling probability and the interview probability increases - firms become more 'picky'. An econometrician not accounting for time-varying selection would measure a fall in the efficiency of matching during the pandemic recession and a negative TFP shock affecting unemployed workers.

The behavior of unemployment implied by the model is shown in Figure 2, which plots the aggregate unemployment rate and the unemployment rates of high-efficiency and low-efficiency workers in the left panel, while the right panel shows the share of the aggregate unemployment rate response that consists of each worker type. Because the steady-state value of $U^l$ is much higher than that of $U^h$, each series is shown as a deviation from its own steady-state value. For type $h$ workers, their unemployment rate rises by less than 3 percentage point above its steady-state value. In contrast, that of type $l$ workers jumps by almost 15 percentage points,

[20] Considering the ratio of total transitions over a quarter relative to the average number employed over the quarter returns a similar increase in transition rates for 2020Q2.

[21] The SPF provides a forecast for unemployment. Using as target variables the 2020Q2 vintage of unemployment and output, the model-implied paths for the $\rho_t^x$ and $D_t$ shocks that are similar to the our baseline and has limited impact on the results. We prefer our specifications, which allows the model to produce a conditional forecast for unemployment – which turns out to be very close to the SPF forecast.

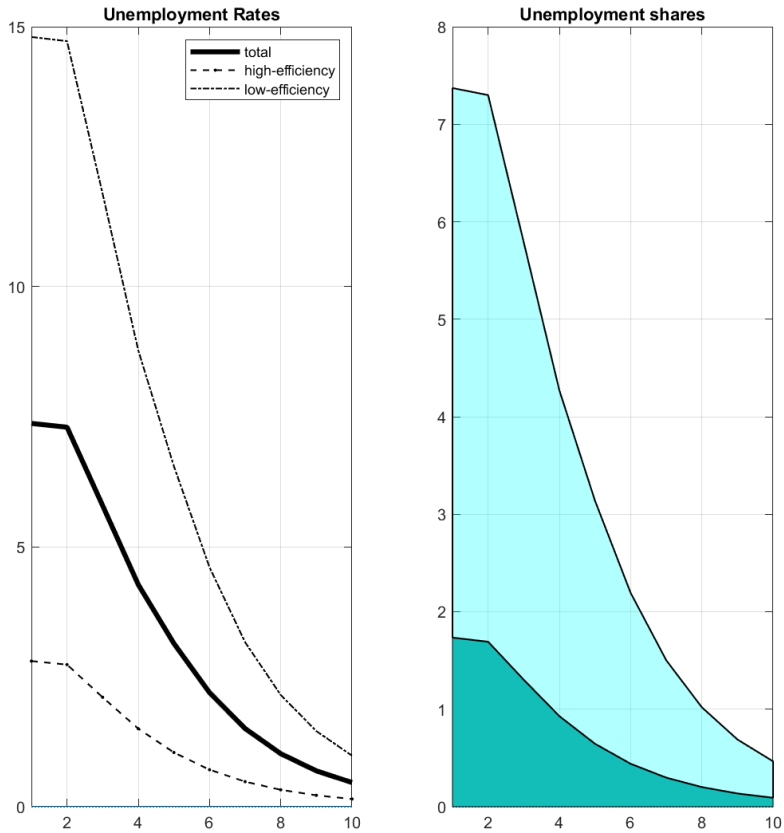


Figure 2: COVID scenario unemployment responses. The left panel shows the responses of $U$, $U^h$, and $U^l$, with each expressed in terms of percentage point deviation from their steady-state values. The right panel shows the total aggregate unemployment rate response and the share accounted for by each worker type: type $l$ (in light) and type $h$ (in dark).

from its steady-state value of just under 10% to almost 25%. Thus, the unemployment rate of type $l$ workers rises more and from a higher steady-state base than the total unemployment rate or the rate for type $h$ workers. And after 6 quarters, the unemployment rate among high-efficiency workers is less than 1 percentage point above its steady-state value, while for low-efficiency workers it is still 4.6 percentage points above steady state. The right panel of the figure shows the composition of total unemployment among worker types. Even though type $l$ workers are less than 40% of the labor force, they account for the bulk of the pandemic-induced higher aggregate unemployment rate.

To understand the reasons for the poor labor market experiences of low-efficiency workers, consider first a non-pandemic recession. Firms reduce their workforce by becoming more selective, retaining fewer low-efficiency workers and screening out more such workers in the interview process. This leads to a rise in $\gamma_t$, the share of low-efficiency workers in the pool of

job seekers which, in turn, reduces the incentive for firms to post job vacancies. In a pandemic scenario, however, social distancing and lockdowns leads to mass layoffs that affect both worker types. Because most workers are high-efficiency types, more type $h$ workers initially flow into the unemployment pool and $\gamma_t$ initially falls, as shown in the top left panel of Figure 3. But the fall in $\gamma_t$ is quickly reversed and then remains persistently above its steady-state value. This reversal reflects greater selectivity by firms; the productivity cutoff value $\bar{a}_t^l$ rises as shown in the top right panel of the figure. Firms become more selective for two reasons. First, the drop in demand for the wholesale good results in a fall in its price relative to the index of sticky retail goods prices; that is, the retail price markup $\mu_t$ rises, as the bottom left panel of the figure shows. The marginal revenue product of a type $l$ worker with idiosyncratic productivity $a_{i,t}^l$ is $a_{i,t}^l \phi^l / \mu_t$. When $\mu_t$ rises, the value of $a_{i,t}^l$ necessary to generate a positive match surplus rises. Second, the value of $\bar{a}_t^l$ also depends on the continuation value of a match. With a persistent rise in $\rho_t^x$ the expected duration of a match and its continuation value falls, implying only very productivity type $l$ workers generate a positive surplus. The rise in $\bar{a}_t^l$ increases the endogenous separation rate, and this further reduces the expected duration of a match. The rate at which type $l$ workers enter the pool of unemployment rises and their exit rate falls, causing $\gamma_t$ to rise above its steady-state value. As both $\gamma_t$ and $\bar{a}_t^l$ increase, the efficiency of the aggregate matching function as defined by $1 - \gamma_t \rho_t^n$ in (7) falls (see bottom right panel).

The responses of job finding rates, vacancy filling rates, vacancies per job seeker, and the average productivity of the pool of unemployed are shown in Figure 4. As the upper left panel shows, the job finding probability falls significantly for type $l$ workers, reflecting the fall in vacancies per job seeker shown in the lower left panel that reduces the probability a worker gets an interview and the fall in the probability a type $l$ worker is hired, conditional on getting an interview. From the perspective of firms, the fall in vacancies per job seeker increases the chances a firm will interview a worker (see dashed line in upper right panel), but the probability of actually successfully hiring rises much less (the solid line in the upper right panel) as firms screen out more of the workers who are interviewed.

Normally, a mass, exogenous destruction of job matches would lead firms to quickly post vacancies to rebuild employment. The recovery of employment is muted in the case of pandemic-induced mass layoffs for two reasons. First, the pandemic involves a negative aggregate demand shock as well as a shock to separation, and this reduces aggregate labor demand. Second, as $\gamma_t$ rises, the average productivity of job seekers falls, as shown in the lower right panel of the figure. This reduces the expected surplus the firm can expect if it posts a new vacancy. It is more likely to interview a low-efficiency worker, and the rise in $\bar{a}_t^l$ means any type $l$ worker the firm does interviews is less likely to be sufficiently productive to generate a positive surplus.

While our discussion has focused on low-efficiency workers as they are the most affected, high-efficiency workers are also adversely affected. The mass layoff due to the separation shock leads to more type $h$ workers in the pool of unemployed. Vacancies per job seeker falls, but once the average productivity of the unemployed begins to fall, this further reduces vacancy

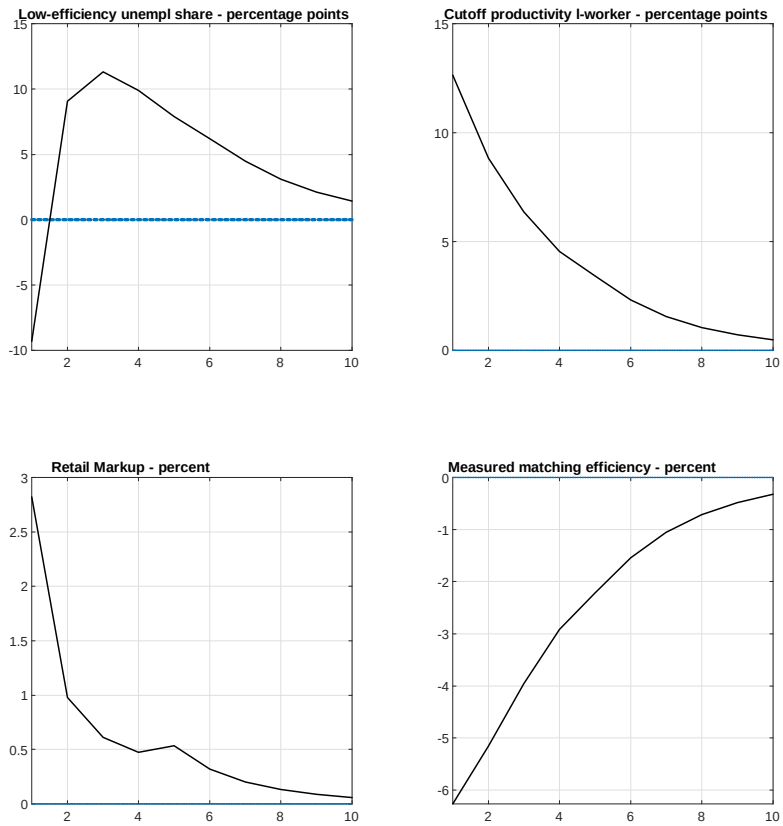


Figure 3: COVID scenario: Responses of $\gamma_t$ (upper left), $\bar{a}_t^l$ (upper right), $\mu_t$ (lower left) and $1-\gamma_t\rho_t^n$ (lower right) to the COVID shocks. Upper panels are percentage point deviations from steady state; lower panels are percent deviations from steady state.

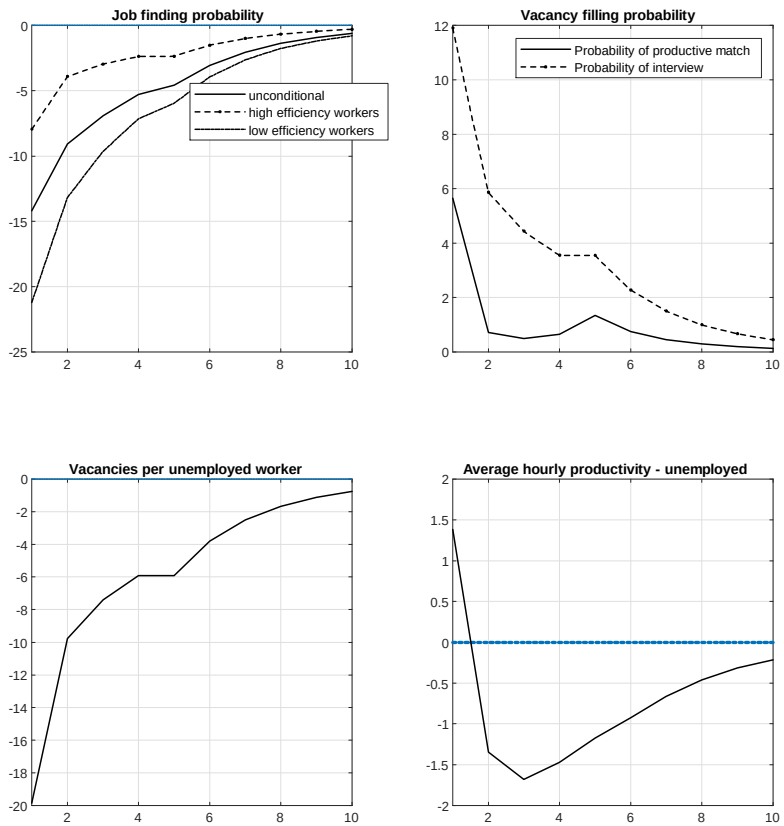


Figure 4: COVID scenario: Responses of job finding probabilities (upper left), vacancy filling probabilities (upper right), vacancies per job seeker (lower left) and average productivity of the unemployed (lower right) to the COVID shocks. Each variable is expressed as a deviation from its steady-state value.

creation. Both developments cause a fall in the probability a type $h$ worker is interviewed. As a result, the job finding rate for these workers declines (see upper left panel). It declines much less than the job finding rate of type $l$ workers, but type $h$ workers do experience longer spells of unemployment as their exit rate declines. The effects of the pandemic shock are exacerbated by the way selection slows the rate at which all workers exit unemployment.

### 3.4 Results on efficiency

Our second major result is that the endogenous responses to the pandemic shock are inefficient and that the resulting inefficiencies disproportionately affects those workers whose average labor market outcomes involve more frequent and longer unemployment spells. The reason is that, as shown in section 2, the surplus generated by such workers in the competitive equilibrium is undervalued relative to its valuation by the social planner. The reverse holds for type

$h$ workers.

To assess the effects of the market distortions, we integrate the wedge between the market and efficient outcomes over the first 10 quarters of the pandemic scenario. Applied to unemployments rates, for example, this would provide a measure of the cumulative excess unemployment after the shocks due to market distortions. These measures of excess unemployment for each worker type are shown in the left panel of Figure 5. Both unemployment rates increase more in the market equilibrium than in the social planner's allocation, but the impact is particularly pronounced for type $l$ workers. After 5 quarters, type $l$ workers have suffered a cumulative efficiency wedge equal to 100% of their steady-state unemployment rate. Since the latter is roughly 10%, they experience an additional 10 percentage points of unemployment over this period that is socially inefficient. In contrast, type $h$ workers, while also experiencing inefficiently high unemployment, suffer cumulative excess unemployment of less than 20% of their steady-state rate of 3%, or approximately an additional 0.6 percentage points of excess unemployment. Thus, the consequences of the pandemic shocks fall disproportionately on type $l$ workers.

With unemployment rates higher and employment lower in the market equilibrium, aggregate output is also inefficiently low in response to the shocks. The cumulative loss in output due to the inefficiency wedge is shown in the right panel of figure 5. After 10 quarters, this efficiency loss totals 4.75% of steady-state output.

To understand the mechanisms generating these efficiency wedges, recall that we have followed the common practice of eliminating steady-state distortions and we have imposed the Hosios condition. Therefore, the only distortions generating any differences between the market and efficient allocations arise from price stickiness and the selection distortion. The efficient equilibrium ensures price stability, so the markup is constant and $\mu_t = 1$. In the market equilibrium, the markup rose as a result of the pandemic shocks. This rise in $\mu_t$ was one of the channels generating a rise in $\bar{a}_t^l$ and an increase in selectivity. Thus, $\bar{a}_t^l$ rises more in the market equilibrium than it would in the efficient allocation; this inefficiency wedge is shown in the top panel of Figure 6. The higher value of $\bar{a}_t^l$ in the market equilibrium implies greater selectivity and a higher endogenous separation rate for low-efficiency workers than would be efficient. As a result, the share of type $l$ workers in the pool of unemployed is also inefficiently high (see middle panel) through both an inefficiently high exit rate from employment and inefficiently low exit rate from unemployment.

The consequences of an inefficiently high values of $\bar{a}_t^l$ and $\gamma_t$ affect the job finding probabilities for both worker types. The composition effect lowers the average productivity of the pool of job seekers, firms have less of an incentive to post vacancies, and the efficiency of the matching function is lower. These developments lead to lower job finding rates for type $l$ workers *and* for type $h$ workers, as shown in the lower left panel of figure 6.

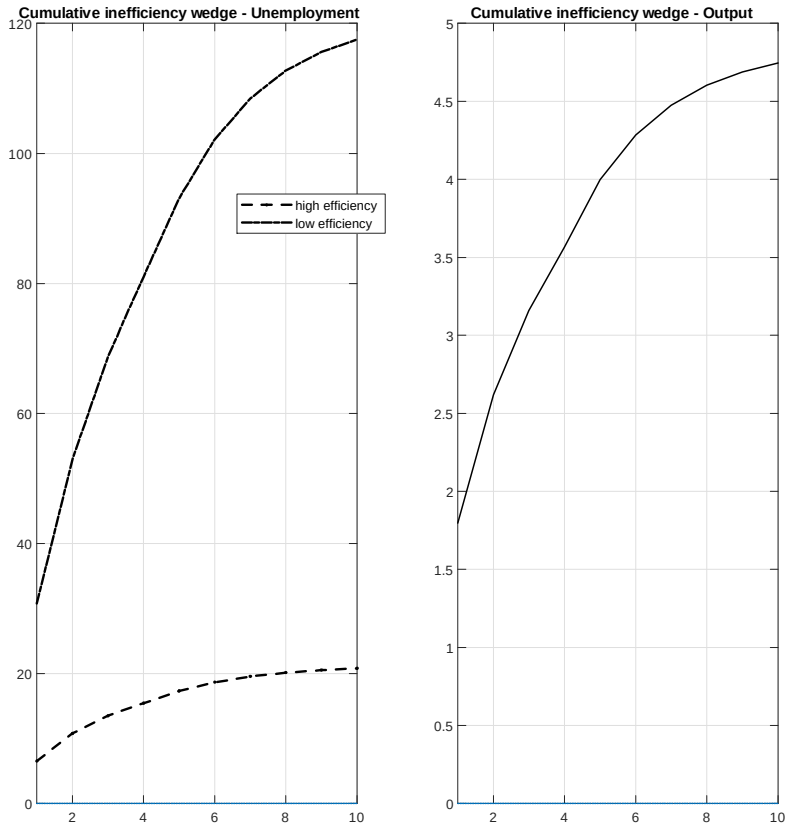


Figure 5: COVID scenario unemployment and output efficiency losses. The cumulative gap between the market outcome and the efficient allocation in response to the COVID shocks for unemployment rates by worker type (left panel) and for output (right panel) measured as a percent of steady-state values.

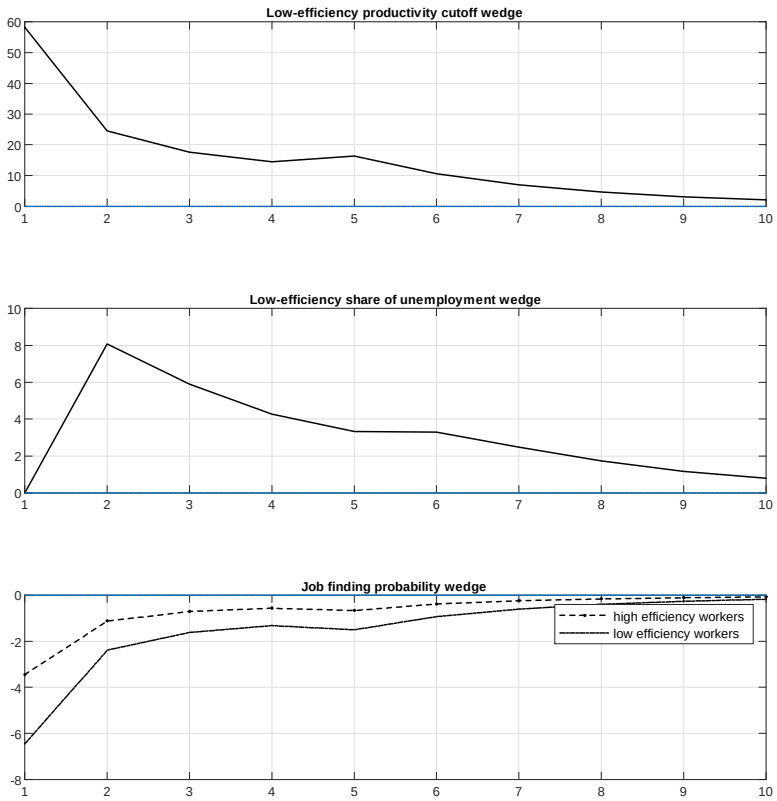


Figure 6: COVID scenario efficiency losses. The difference between market and efficient responses to the COVID shocks for $\bar{a}_t^l$ (upper panel), $\gamma_t^l$ (middle right), and job finding rates $k_t^w$ and $k_t^{w,l}$ (lower left) measured as a percent of steady-state values.

# 4 Temporary layoffs

Our baseline model assumes that all exogenous separations are permanent separations. Workers flowing into unemployment directly enter the pool of searching workers. However, in the initial phase of the COVID-19 recessions, a large fraction of separations consisted of temporary layoffs, as Barrero et al. (2020) and Kudlyak and Wolcott (2020) have emphasized. Figure 7 shows the atypical behavior of temporary layoffs in early 2020. According to the BLS household survey, the share of job losers on temporary layoff averaged 12.4% between January 2000 and December 2019 before spiking at 77.9% in April 2020. Figure 8 shows the same series as the previous figure but focuses in on the period from December 2019 to December 2020. After peaking in April, temporary layoffs declined steadily until November before rising slightly in December as COVID-19 cases again spiked in the U.S. Based on previous recessions, Barrero et al. (2020) estimated that 42% of recent layoffs would result in permanent job losses. It is therefore important to see how dynamics are affected when a large share of layoffs are, at least initially, temporary. If these workers are recalled to their former jobs as the economy begins to reopen, the speed at which employment recovers may be much faster than predicted by models that ignore temporary layoffs.[22] In this section, therefore, we extend our model to include both permanent and temporary layoffs.

It is useful to note first how unusual the COVID-19 behavior of temporary layoffs was. Over the 420 months between January 1985 and December 2019, a period that includes the Great Moderation, the Global Financial Crisis, and the post-financial crisis recovery, the covariance between the share of temporary layoffs and the unemployment rate was $-1.63$, indicating that the share of workers on temporary layoff was procyclical. Adding the twelve months January to December 2020 to the sample causes this covariance to flip to a positive 2.09.[23] The COVID-19 surge in both unemployment and in the share of those on temporary layoff was unprecedented.

To allow for temporary separations, we assume a fraction $0 \leq \Gamma_t \leq 1$ of exogenous separations at time $t$ are temporary. Workers on temporary layoff are assumed to be recalled at a constant rate $0 < r \leq 1$. All recalled workers of type $h$ are rehired. However, type $l$ workers receive idiosyncratic productivity realizations, and consequently, the productivity of a type $l$ worker when recalled may fail to exceed $\bar{a}_t^l$. Thus, only a share $(1 - \rho_t^n)$ of type $l$ workers who are recalled at time $t$ are actually rehired. We assume those who are screened out after recall enter the pool of permanently separated job seekers.

If $\Gamma_t$ were constant, our model would predict that the share of unemployed on temporary layoff would fall in a recession, consistent with the pre-pandemic evidence, as total separations include both exogenous and endogenous separations and the latter rise during a recession. To capture the positive co-movement seen during the pandemic, we assume that the positive shock

[22] Some countries have indeed observed fast rebounds: Norway saw its unemployment rate climb from 2% to 11% within 60 days after its first COVID cases, and then fall back to 7% after a further 30 days as the economy started reopening. However, many separations are expected to become permanent, as future pandemic prevention policies may call for structural reductions in some sectors, such as travel and hospitality.

[23] The correlation coefficients are $-0.61$ and 0.21 for the two periods.

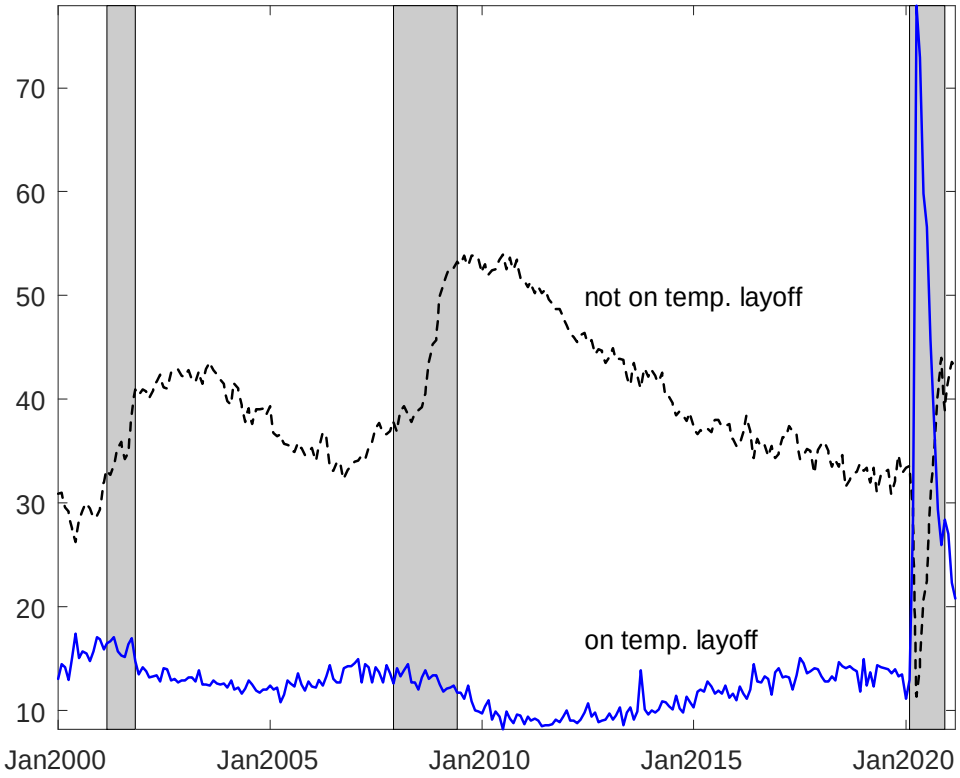


Figure 7: Share of job lossers on temporary layoffs and not on temporary layoffs. Source: CPS series LNS13023654 and LNS13026511. Business cycle recessions.shown by shaded regions (NBER dating).

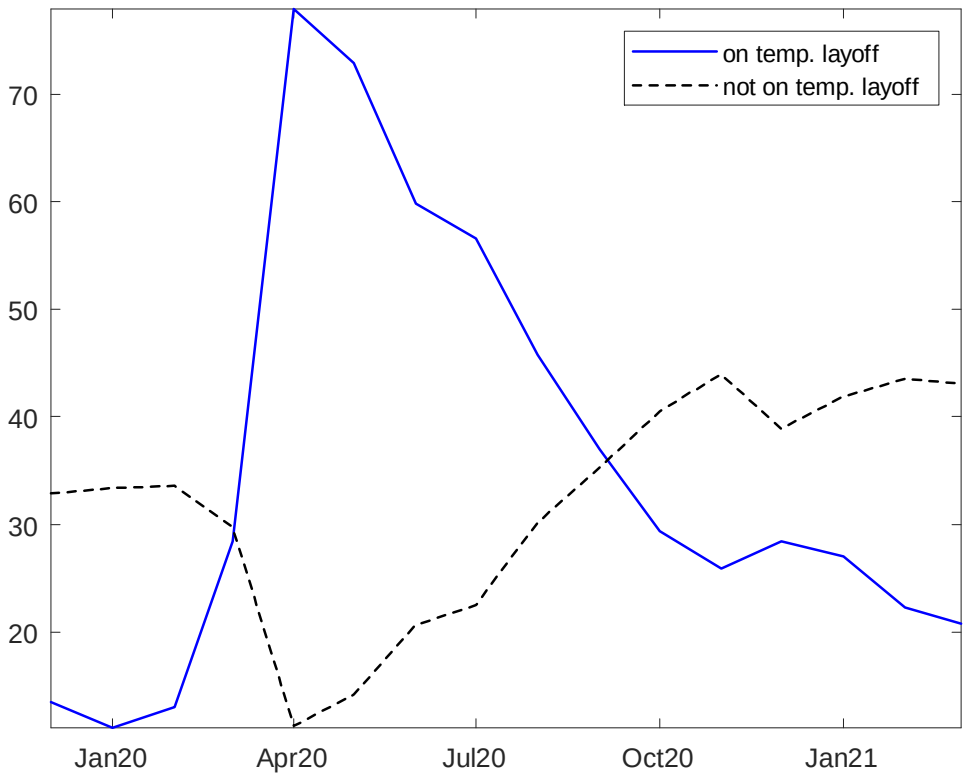


Figure 8: Share of job lossers on temporary layoffs and not on temporary layoffs. Source: CPS series LNS13023654 and LNS13026511.

to total exogenous separations is accompanied by a positive shock to $\Gamma_t$.

The stock of workers of each type that are on temporary layoff evolves according to

$$T_t^j = (1-r)\left(T_{t-1}^j + \Gamma_t \rho_t^x N_{t-1}^j\right), \tag{15}$$

for $j = h,\ j$, where $T_t^j$ is equal to the number of type $j$ workers on temporary layoff. Note we assume that some workers put on temporary layoff in the current quarter may be recalled within the quarter.

The number of job seekers, $S_t$, equals the total number of workers unmatched at the start of the period, $1 - N_{t-1}$, plus those who do not survive the exogenous separation hazard, $\rho_t^x N_{t-1}$, minus those previously separated workers put on temporary layoff awaiting recall to their previous job. Letting $T_t = T_t^h + T_t^l$,

$$S_t = 1 - (1-\rho_t^x)\, N_{t-1} - T_t.$$

Total new matches equal new hires plus recall hires, or

$$H_t = (1-\gamma_t \rho_t^n)\, k_t^w S_t + r\left[T_{t-1}^h + \rho_t^x \Gamma_t N_{t-1}^h + (1-\rho_t^n)\left(T_{t-1}^l + \rho_t^x \Gamma_t N_{t-1}^l\right)\right].$$

Employment in period $t$ consists of surviving matches plus new matches:

$$\begin{aligned} N_t \;=\; & (1-\rho_t^x)\left[\left(1-\xi_{t-1}\right) + (1-\rho_t^n)\,\xi_{t-1}\right] N_{t-1} + (1-\gamma_t \rho_t^n)\, k_t^w S_t \\ & + r\left[T_t^h + \rho_t^x \Gamma_t \left(1-\xi_{t-1}\right) N_{t-1}\right] + r\left(1-\rho_t^{n,l}\right)\left(T_t^l + \rho_t^x \Gamma_t \xi_{t-1} N_{t-1}\right). \end{aligned} \tag{16}$$

Finally, the share $\xi_t$ of type $l$ workers among the employed is given by:

$$\xi_t = (1-\rho_t^n)\left[\frac{\xi_{t-1}\,(1-\rho_t^x)\, N_{t-1} + \gamma_t k_t^w S_t + r\left(T_{t-1}^l + \rho_t^x \Gamma_t \xi_{t-1} N_{t-1}\right)}{N_t}\right],$$

where the last two terms in the numerator consist of those type $l$ who are interviewed and not screened out, $(1-\rho_t^n)\,\gamma_t k_t^w S_t$, and those recalled but not screened out, $(1-\rho_t^n)\, r\left(T_{t-1}^l + \rho_t^x \Gamma_t \xi_{t-1} N_{t-1}\right)$.

## 4.1 Calibration

Our model of temporary layoffs adds two new parameters to the model: the steady-state value $\Gamma$ and the recall rate $r$. To allow the introduction of temporary separations to have the strongest possible impact on the dynamics of the model, we assume that all workers flowing into the pool of temporary layoffs eventually get recalled for a job interview without the need for a vacancy being posted, and that the probability of a recall interview within four quarters of the initial separation is 95%. This implies a quarterly recall hazard $r$ of 53%. In steady-state, over 99% of recalled workers enter into a productive match, while the recall share of workers finding a match endogenously falls after the start of the COVID-19 recession since selected

separations increase sharply.

We parameterize the steady-state share of workers in temporary unemployment relative to the total stock of unemployed to 13%, a value in line with the share reported in Kudlyak and Wolcott (2020) for the 1985-2019 sample. This share, together with the quarterly recall rate, implies that the steady-state share of exogenous separations flowing into the pool of temporary unemployment is equal to 6.44%.

Next we parameterize the shock $\Gamma_t$ to target the path for the share of workers on temporary unemployment relative to the total stock of unemployed $T_t/U_t$ reported from the BLS for the first four quarters of the COVID-19 recession. In the first quarter the targeted share is equal to 70%, approximately equal to the average of the temporary unemployment share over the months of March, April, May 2020. The temporary unemployment share declines in the second, third and fourth quarter of the recession to 46.5%, 27.9% and 23.4%..

### 4.2 A COVID-19 recession with temporary layoffs

We now discuss how temporary layoffs affect the dynamic adjustment of unemployment and its distribution across worker types relative to the model of section 3 in which all separations were permanent. Our main result is that the use of temporary layoffs primarily benefits high-efficiency workers. Such workers are more likely to return directly to employment without participating in the search and matching process, and they experience less overall unemployment. Even though the recall rate is the same for both worker types, some recalled type $l$ workers are screened out and not rehired. This leads to a larger rise in the share of low-efficiency workers among the pool of job seekers relative to the benchmark case with only permanent layoffs. Because firms can rebuild matches through recalls, fewer vacancies are posted and firms become more selective.

These implications of allowing for temporary layoffs are shown in Figure **9**. All panels show differences in the dynamics of a variable relative to the case discussed in section 3 where layoffs were permanent. The upper left panel shows that the number of searching workers falls relative to the benchmark permanent layoff model as some non-matched workers are on temporary layoff awaiting recall and are not searching. This affects type $h$ workers primarily; type $l$ workers see a smaller reduction in active job searchers as employed workers with sufficiently low productivity outcomes separate permanently and will not end up in the temporary layoff pool. Unemployment, shown in the upper right panel, is lower in the temporary layoff model as some workers on temporary layoff can be recalled within the same quarter. Almost all the reduction comes from lower unemployment among type $h$ workers as some low-efficiency workers are recalled but not rehired.

Vacancies per job seeker falls with temporary layoffs, but as the lower left panel shows, labor market tightness $\theta$ falls more with temporary layoffs. Firms now have access to recalled workers, reducing the need to hire in the search market. When they do post a vacancy, they are also more likely to hire a low-efficiency worker, an outcome that also dampens the incentive

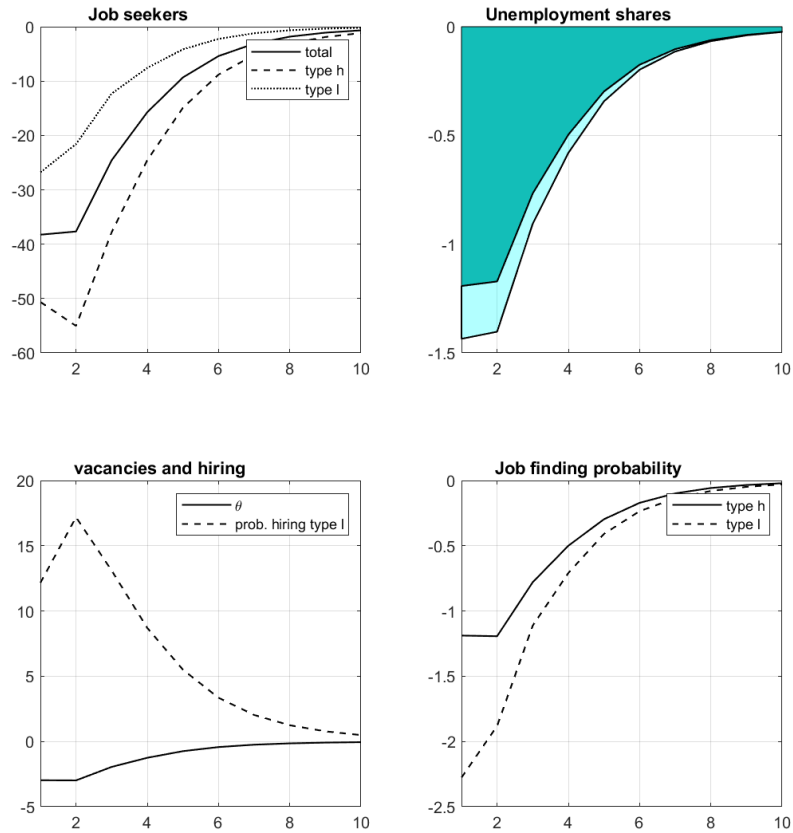


Figure 9: Temporary layoff model: Each panel shows the difference between the selected variable in the model with temporary layoffs and in the benchmark model in which all layoffs are permament. Panel 2 shows the total fall in unemployment in percentage points divided into the share accounted for by each worker type shown as the difference between the temporary layoff model and the permanent layoff model: type $l$ (light), type $h$ (dark). Other panels expressed as percent of steady state.

to post vacancies. Finally, the probability of finding a job in the search and matching market falls for both high- and low-efficiency workers, but the lower right panel of the figure shows that this probability falls more in the presences of temporary layoffs, and it falls more for low-efficiency workers.

The composition of the pool of job searchers measured by $\gamma_t$ and the cutoff productivity level $\bar{a}_t^l$ are again key to understanding the effects on the different worker types. The share of type $l$ workers in the unemployment pool is higher in the temporary layoff model as more type $h$ workers are able to move directly from temporary layoff to employment through recall. The greater increase in $\gamma_t$ in the presence of temporary layoffs for the increase probability a firm with a vacancy will fill it would a type $l$ workers (see figure 9). Firms interview more type $l$ workers but they also become more selective, as $\bar{a}_t^l$ rises, reflecting the larger increase in

the markup that occurs with temporary layoffs. The markup rises relative to the permanent layoff model because with unemployment lower in the temporary layoff model, employment is higher, wholesale output is higher, and the price of the wholesale good falls more relative to the sticky retail price index.[24]

Finally, the introduction of temporary layoffs lowers the unemployment rate for both worker types. However, when a large fraction of the jobs lost as a result of mass layoffs take the form of temporary layoffs, as they did in the early days of the COVID-19 recession, the composition of unemployment is strongly skewed towards low-efficiency workers. One way to measure this skewness is to calculate the ratio of $U^l$ to $U^h$. With only permanent layoffs, the ratio $U^l/U^h$ is equal to 3.25 on impact, 3.6 after one year, and 4.03 after two years. Allowing for temporary layoffs, this ratio rises to 9.78, 7.3 and 5.7 at the same three horizons. In this sense, temporary layoffs skew the burden of the recession towards low-efficiency workers.

# 5 Monetary policy

In this section, we use our temporary layoff model to investigate how different worker types, and the aggregate economy, are affected by alternative monetary policy rules in the face of the COVID-19 shock. The model provides a framework to discuss the impact of policies on inequality across workers and the implications of using policies explicitly aimed at being more inclusive in supporting workers across the productivity and wage distribution.

In previous sections, our baseline policy rule given in (10) assumed the nominal interest rate reacted only to inflation. We now consider modifications to this rule that are designed to more directly target labor market developments. Specifically, we consider rules of the form

$$\ln(1 + i_t) = -\ln\beta + \omega_\pi \pi_t + \omega_x x_t, \tag{17}$$

where $x_t$ is a labor market variable. We consider three alternatives for $x_t$: the aggregate unemployment rate $U_t$, the total separation rate $\rho_t$, and the unemployment rate of type $l$ workers $U_t^l$. Because cyclical fluctuations in separations primarily affect type $l$ workers, reacting to $\rho_t$ is potentially a means of targeting the group that experiences more frequent and longer spells of unemployment.[25] $U_t^l$ is a direct measure of the impact of the recession on type $l$ workers, but, within the context of the model, it is unobservable. Analyzing a rule that responds to $U_t^l$ provides a 'best case' scenario in terms of what monetary policy might achieve in reducing employment fluctuations among type $l$ workers. For comparison, we also consider a policy that maintains price stability.

As in our baseline rule, we set $\omega_\pi = 1.5$. The response coefficient on the labor market variable, $\omega_x$, is set to achieve a target stabilization goal. For the $U_t$ and $\rho_t$ rules, we assume

[24] Wholesale output rises because employment rises but also because the increase in $\bar{a}_t^l$ means the average productivity of type $l$ workers who are retained or hired is higher.

[25] Berger et al. (2019) examined policy rules targeting the layoff rate, and found that they approximate well the behaviour of the Federal Reserve and produced good outcomes.

the goal is to lower the first period recessionary impact of the COVID-19 shock on aggregate unemployment by 25%; $\omega_x$ is set to achieve this goal. This policy also achieves approximately a 25% reduction in the impact of the pandemic shock on $U_t^l$. For the $U_t^l$ rule, we set a more ambitious goal and assume the monetary authority specifically aims to reduce the first period pandemic impact on $U_t^l$ by 50%.

We use three metrics to evaluate these alternative policies. The first is the output loss, measured as the cumulative loss in output, expressed as a percentage of steady-state output, over the first two years of the pandemic. The second metric we call the 'type $l$ unemployment loss'. The time-$T$ loss is defined as the cumulative excess percentage points of total unemployment $U_t$ accounted for by $l$-workers over the horizon $[0 : T]$.[26] We measure the loss over the first two years of the pandemic. We also calculate the corresponding type $h$ unemployment loss. Our third metric, called the inequality ratio, is simply the ratio of the type $l$ unemployment loss to the type $h$ unemployment loss. A smaller value of this ratio implies that a policy results in less dispersion in the behavior of unemployment across the two worker types. In the context of the COVID-19 recession, which affects disproportionately type $l$ workers, a lower inequality ratio implies a rebalancing of the burden of the recession towards a more equal distribution across worker groups. Finally, we compute a symmetric *sacrifice ratio* defined as the cumulative absolute deviation of inflation from its target divided by the cumulative reduction in unemployment from its peak in the first period of the pandemic recession. Using a two year horizon, this measure equals

$$Sacrifice\ Ratio_T = -\sum_{t=1}^{T=8} |\pi_t - \pi_{ss}| / \sum_{t=1}^{T=8} (U_t - U_1).$$

It provides an assessment of the cost in terms of inflation for each percentage point of fall in unemployment from the peak that would occur conditional on the baseline policy.

## 5.1 Comparing alternative rules

The implications of the baseline rule and the alternative rules are reported in columns 1 to 4 of Table 2. Comparing columns 1 and 2 shows that the rule calibrated to reduce the rise in $U_t$, not surprisingly, also reduces the output loss. It also reduces the unemployment loss for both worker types. Type $l$ workers benefit greatly from this policy, with their cumulated unemployment loss falling from 24.56% to 17.09%, a 30% reduction. In comparison to the baseline policy, the $U_t$ policy succeeds in reducing the unemployment cost for type $l$ workers relative to type $h$ workers, as measured by the inequality ratio, which falls from 8.07 to 6.88, a 17% decline. Note that the fall in the inequality ratio is less than proportional to the fall in the unemployment loss for low-efficiency workers, reflecting the fact that this policy also reduces

[26] At each $t$ the excess $l-$unemployment rate is computed as $\alpha U_t - U_{ss}$ where $\alpha = (N_t^l - N_{ss}^l)/(N_t - N_{ss})$.Scaling the number of unemployment workers by the total number of unemployed allows us to compare the unemployment losses across workers types.

the unemployment loss for high-efficiency workers. By benefiting all workers, stabilizing $U_t$ has a more modest impact on how the burden of the recession is allocated across the two types of workers. Finally, responding to $U_t$ is also effective in reducing the sacrifice ratio, as seen in the table's bottom row.

Column 3 in Table 2 reports the results if monetary policy responds to total separations. In terms of the measures of loss, this policy is similar to one that just responds to aggregate unemployment. Both the $U_t$ rule and the $\rho_t$ rule target the same reduction in the impact of the pandemic recession on aggregate unemployment, so the result that they achieve similar output losses is not surprising. The results from columns 2 - 3 indicate that by targeting aggregate labor market outcomes, monetary policy can reduce the aggregate output loss and unemployment losses for all workers. Policy has some effect in reallocating the impact from a pandemic shock across worker types, and thus reducing the inequality of the recession burden, but this effect is more modest. The similarity of the $l$-unemployment loss and $h$-unemployment losses under the $U_t$ and $\rho_t$ rules indicates the measures of unemployment loss move roughly in line with the aggregate economy, regardless of the specific variable added to the policy rule.

The general explanation for these results is that monetary policy affects the aggregate economy but cannot ensure that expansionary policy designed to dampen the adverse impacts of the pandemic will propagate in a manner that helps those workers most affected by the recession. Focusing on the behavior of the markup helps understand why. Monetary policy is able to reduce unemployment by reducing interest rates and increasing aggregate demand. This increased demand for retail goods also boosts the demand for wholesale goods and the demand for labor. Under the baseline policy, the retail price markup rises as the recessionary drop in demand causes wholesale prices to fall, as shown in the top panel of Figure 10. All the alternative policy are more expansionary at the onset of the pandemic. As such, they support aggregate demand and output. This limits the rise in the markup, and, as seen in the bottom panel, the rise in $\bar{a}_t^l$, helping to reduce the inflow of type $l$ workers into unemployment and increase their exit rate from unemployment. However, the effect on the inequality ratio is limited because the markup affects the marginal revenue product of both worker types; the value of the output produced by a type $h$ worker is $\phi^h/\mu_t$ and that by a type $l$ worker is $a_{i,t}^l\phi^l/\mu_t$. To the extent $\mu_t$ rises less, the match surpluses for low-efficiency workers *and* high-efficiency workers fall less. Both types benefit, but because both worker types benefit, the ability of monetary policy to have a significant and differential impact on the different worker groups is limited.

The inverse of $\mu_t$ is real marginal cost for retail firms. By limiting the rise in $\mu_t$, the policies stabilizing $U_t$ also stabilize inflation. As row (5) of the table shows, the policies responding to $U_t$ (col. 2) or $\rho_t$ (col. 3) each has lower sacrifice ratio than the baseline rule that responds only to inflation (col. 1).

Temporary layoff model with alternative policies rules. Panels show the responses of $\mu_t$ (top) and $\bar{a}_t^l$ (bottom) to the pandemic shocks under three different rules; the baseline rule (solid line), the $U_t$

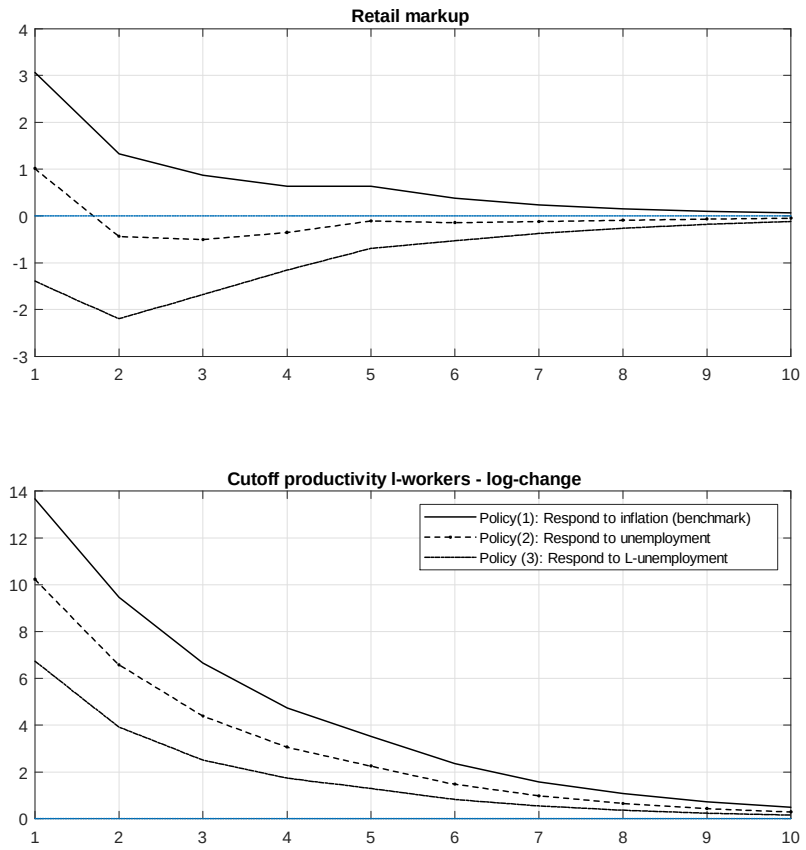


Figure 10: Temporary layoff model with alternative policies rules. Panels show the responses of $\mu_t$ (top) and $\bar{a}_t^l$ (bottom) to the pandemic shocks under three different rules; the baseline rule (solid line), the $U_t$ rule (dashed line), and the $U_t^l$ rule (dotted line)

rule (dashed line), and the $U_t^l$ rule (dotted line)

Column 4 of Table 2 shows the outcome if monetary policy could react directly to the labor market conditions of type $l$ workers by responding to $U_t^l$. This policy is the most successful at reducing the inequality ratio. This policy delivers a 50% reduction of in the unemployment rate of type $l$ workers at the onset of the recession and produces an approximately equal fall in the aggregate unemployment rate.[27] Table 2 shows that this policy will also considerably lower the unemployment inequality ratio. Figure 10 also shows that the policy will be more inclusive; limiting the rise $U^l$ is sufficiently expansionary that the markup actually falls. As a result, the rise in the cutoff productivity level $\overline{a}_{i,t}$ is reduced by about 50% relative to the benchmark policy. With $\bar{a}_t^l$ rising less, the endogenous separation rises less, implying fewer type $l$ workers are screened out and more who are in matches are retained. However, this improvement in the labor market outcomes for type $l$ workers comes at a considerable cost in terms of inflation volatility. The protracted fall in $\mu_t$ below its steady state leads to higher inflation and to a worse sacrifice ratio, which rises from 0.18 under the baseline policy to 0.25 as the sacrifice ratio punishes the absolute deviations of inflation from steady state.

A comparison of columns 2 and 3 for the $U_t$ and $\rho_t$ policies with column 4 provides a measure of what policy might achieve in targeting those workers who have poorer labor market outcomes. Because the $U_t^l$ policy targets a more aggressive stabilization of the unemployment rate of low-efficiency workers, it leads to a much more expansionary policy response to the pandemic. The policy does succeed in reducing the output loss and achieves a large reduction in the cumulative unemployment loss for the targeted group of workers. It also achieves some reduction in the inequality ratio. However this comes at a the cost of a significant rise in the sacrifice ratio.

Finally, column (5) of Table 2 reports results for policy of price stability. It results in very similar real outcomes outcome to a policy reacting to the endogenous separation rate while (by construction), driving the sacrifice ratio to 0. The table shows that the outcomes in terms of output and type-specific unemployment loss, and in terms of inequality, are also very close. As is common in new-Keynesian models with forward-looking price setting, the policymaker's commitment to a price level target affects firms' expectations - and in our model also the surpluses' continuation values - leading to more favorable outcomes, relative to our baseline rule that allows inflation to fluctuate.

Overall, our analysis finds that the monetary authority can reduce the impact of the COVID-19 shock on the aggregate economy and on the adverse employment outcomes of low-efficiency workers. However monetary policy is not very effective in reallocating the impact of a pandemic shock across worker types. Reducing the burden of the recession for the low-efficiency group of workers will necessarily lead to also supporting employment for high-efficiency workers. A policy that could directly target the unemployment rate of a specific

[27] The policymaker can directly observe only proxy-measures for the level of unemployment $U^l$. We give the best chance to the policy by assuming that it can directly react to $U^l$.

worker group is more successful in reducing our inequality ratio, but it comes at a high cost in terms of inflation volatility. This does not imply that monetary policy is ineffective – rather, the monetary policy transmission channel is not adequate to target specific groups of workers affected by the pandemic. Surprisingly, a policy of price stability performs very closely to the more employment 'inclusive' policies we consider in terms of labor outcomes across the worker types, while it also eliminates inflation volatility.

# 6 Conclusions

We have used a calibrated new Keynesian model with labor heterogeneity and selection in a search and matching labor market to investigate the distribution of unemployment across workers in a pandemic recession such as that caused by COVID-19. We model the pandemic as a negative demand shock and a spike in mass layoffs and use our model to compare the market adjustment to a COVID shock to the adjustment in the social planner's allocation. We identify a new externality that arises when individual firms ignore the effects their separation and hiring decisions have on the composition of the pool of unemployed workers. During a pandemic-induced recession, the resulting distortion acts to disproportionately worsen the labor force outcomes for those workers who on average have more frequent and longer spells of unemployment. Inefficient selection in endogenous hiring and retention decisions adversely affects the quality of the unemployment pool, leading firms to post fewer vacancies. All workers end up experiencing poorer labor market outcomes. These results continue to hold even if a large fraction of the layoffs take the form of temporary layoffs, as they did early in the COVID-19 recession.

COVID-19 had heterogeneous effects across many economics dimensions, with disparate effect across sectors, industries, regions, and individuals. We have focused on heterogeneity across workers and their employment experiences because of the interest expressed by central banks, particularly in the U.S., over the distributional impact of monetary policy on workers who suffer high rates of unemployment. While our model suggests that by responding to labor market variables monetary policy can help stabilize unemployment in the face of recessionary shocks, thereby benefiting all workers, the ability to affect the distribution of unemployment across worker types is more limited. Monetary policy can reduce the inequality of the recessionary burden falling on workers with different productivity levels, but only at a considerable cost in terms of inflation volatility. Because the markup plays a central role in determining the value of job matches and therefore the endogenous separation rate, a policy of price stability performs very similarly to a policy rule that responds to total employment separations. To significantly affect the disproportionate unemployment burden that falls on some workers, our results suggest tools other than monetary policy, such as taxes and/or subsidies, may be more effective instruments for dealing with cyclical variations in impact of unemployment across workers.

## References


Ahn, H. J. and J. D. Hamilton (2019). Heterogeneity and unemployment dynamics. *Journal of Business and Economic Statistics*, 1–16.

Aum, S., S. Y. Kee, and Y. Shin (2020). COVID-19 Doesn't Need Lockdown to Destroy Job: The Effect of Local Outbreaks in Korea. *NBER Working Paper No. 27264*.

Baker, G. P. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy 100*(3), 598–614.

Baley, I., A. Figueiredo, and R. Ulbricht (2020). Mismatch Cycles. *SSRN Electronic Journal*.

Baqaee, D. R. and E. Farhi (2020). Supply and Demand in Disaggregated Keynesian Economies with an Application to the Covid-19 Crisis.

Barrero, J. M., N. Bloom, and S. J. Davis (2020). COVID-19 is also a reallocation shock. *NBER Working Paper No. 27137*.

Berger, D., I. Dew-Becker, L. D. W. Schmidt, and Y. Takahashi (2019). Layoff Risk, the Welfare Cost of Business Cycles, and Monetary Policy. *Working Paper*.

Bergman, N., D. A. Matsa, and M. Weber (2021). Heterogeneous Labor Market Effects of Monetary Policy. *SSRN Electronic Journal*.

Bernstein, J., A. W. Richter, and N. A. Throckmorton (2020). COVID-19: A View from the Labor Market. *Federal Reserve Bank of Dallas, Working Paper No. 2010*.

Bils, M., Y. Chang, and S.-B. Kim (2012). Comparative advantage and unemployment. *Journal of Monetary Economics 59*(2), 150–165.

Cheng, W., P. Carlin, J. Carroll, S. Gupta, F. L. Rojas, L. Montenovo, T. D. Nguyen, I. M. Schmutte, O. Scrivner, K. I. Simon, C. Wing, and B. Weinberg (2020). Back to Business and (Re)Employing Workers? Labor Market Activity During State COVID-19 Reopenings. *NBER Working Paper No. 27419*.

Cortes, G. M. and E. Forsythe (2020a). Impacts of the Covid-19 Pandemic and the CARES Act on Earnings and Inequality. pp. 1–42.

Cortes, G. M. and E. Forsythe (2020b). The Heterogeneous Labor Market Impacts of the Covid-19 Pandemic.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996). *Job Creation and Job Destruction*. Cambridge, MA: The MIT Press.

Dickens, W. T. and R. K. Triest (2012, jan). Potential Effects of the Great Recession on the U.S. Labor Market. *The B.E. Journal of Macroeconomics 12*(3).

Fornaro, L. and M. Wolf (2020). Covid-19 Coronavirus and Macroeconomic Policy: Some Analytical Notes. *Working Paper* (March), 1–8.

Gertler, M., L. Sala, and A. Trigari (2008). An estimated monetary DSGE model with unemployment and staggered Nash wage bargaining. *Journal of Money, Credit and Banking 40*(8), 1713–1764.

Gregory, V., G. Menzio, and D. Wiczer (2020). Pandemic Recessions: L- or V-Shaped. *FRB Minneapolis Quarterly Review* (May).

Gregory, V., G. Menzio, and D. G. Wiczer (2021). The Alpha Beta Gamma of the Labor Market. *NBER Working Paper No. 28663*.

Grigsby, J. (2020). Skill Heterogeneity and Aggregate Labor Market. *Working paper*, 0–83.

Guerrieri, V. (2007). Heterogeneity and unemployment volatility. *Scandinavian Journal of Economics 109*(4), 6667–693.

Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2020). Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages? *NBER Working Paper No. 26918*.

Hall, R. E. (2015). Quantifying the Lasting Harm to the U. S. Economy from the Financial Crisis. In J. A. Parker and M. Woodford (Eds.), *NBER Macroeconomics Annual 2014*, Volume 29, pp. 71–128. University of Chicago Press.

Hall, R. E. (2017). High discounts and high unemployment. *American Economic Review 107*(2), 305–330.

Hall, R. E. and S. Schulhofer-Wohl (2018). Measuring job-finding rates and matching efficiency with heterogeneous job-seekers. *American Economic Journal: Macroeconomics 10*(1), 1–32.

Hornstein, A., P. Krusell, and G. L. Violante (2011). Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review 101*(7), 2873–2898.

Jackson, P. and V. Ortego-Marti (2020). Skill Loss during Unemployment and the Scarring Effects of the COVID-19 Pandemic. *Working Paper*.

Kahn, L. B., F. Lange, and D. G. Wiczer (2020). Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims. *NBER Working Papers*.

Kapicke, M. and P. Rupert (2020). Labor Markets during Pandemics.

Kocherlakota, N. R. (2020). 21st century macro. *NBER Working Paper No. 26791*.

Kospentaris, I. (2020). Unobserved Heterogeneity and Skill Loss in a Structural Model of Duration Dependence. *Review of Economic Dynamics 39*, 280–303.

Kudlyak, M. and E. Wolcott (2020). Pandemic Layoffs. (May), 1–21.

Leduc, S. and Z. Liu (2020). The weak job recovery in a macro model of search and recruiting intensity. *American Economic Journal: Macroeconomics 12*(1), 310–343.

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

Nagypal, E. (2007). Learning-by-doing versus learning about match quality. *Review of Economic Studies 74*(2), 537–566.

Nagypal, E. and D. T. Mortensen (2007). Labor-market volatility in matching models with endogenous separations. *Scandinavian Journal of Economics 109*(4), 645–665.

Powell, J. H. (2020). New economic challenges and the Fed's monetary policy review. *Jackson Hole Symposium*.

Pries, M. J. (2008). Worker heterogeneity and labor market volatility in matching models. *Review of Economic Dynamics 11*, 644–687.

Ravenna, F. and C. E. Walsh (2008). Vacancies, unemployment, and the Phillips curve. *European Economic Review 52*(8), 1494–1521.

Ravenna, F. and C. E. Walsh (2011). Welfare-Based Optimal Monetary Policy with Unemployment and Sticky Prices: a Linear-Quadratic Framework. *American Economic Journal: Macroeconomics 3*(April), 130–162.

Ravenna, F. and C. E. Walsh (2012). Screening and Labor Market Flows in a Model with Heterogeneous Workers. *Journal of Money, Credit and Banking 44*(2), 31–71.

Ravn, M. O. and V. Sterk (2017). Job uncertainty and deep recessions. *Journal of Monetary Economics 90*, 125–141.

Sala, L., U. Söderstrom, and A. Trigari (2008, jul). Monetary Policy under Uncertainty in an Estimated Model with Labor Market Frictions. *Journal of Monetary Economics 55*(5), 983–1006.

Thomas, C. (2008). Search and Matching Frictions and Optimal Monetary Policy. *Journal of Monetary Economics 55*, 936–956.

Trigari, A. (2009, feb). Equilibrium Unemployment, Job Flows, and Inflation Dynamics. *Journal of Money, Credit and Banking 41*(1), 1–33.

Walsh, C. E. (2003). Labor Market Search and Monetary Shocks. In S. AltuÄİ, J. Chadha, and C. Nolan (Eds.), *Elements of Dynamic Macroeconomic Analysis*, pp. 451–486. Cambridge, U.K.: Cambridge Univeristy Press.

Walsh, C. E. (2005). Labor market search, sticky prices, and interest rate policies. *Review of Economic Dynamics 8*(4), 829–849.

| Table 1: Parameters and Steady State Values | | |
|---|---|---|
| Targeted Steady State Values | | |
| Unemployment rate | $u_{ss}$ | 5.6% |
| Unemployment rate: $l-efficiency$ labor | $u^l_{ss}$ | 9.87% |
| Unemployment rate: $h-efficiency$ labor | $u^h_{ss}$ | 2.97% |
| Average hours per worker | $h^{av}_{ss}$ | 0.33 |
| Vacancy posting cost share of output | $\frac{\kappa V_{ss}}{Y_{ss}}$ | 0.015 |
| Probability of vacancy matched with applicant | $k^f_{ss}$ | 0.9 |
| Implied Parameters and Steady State Values | | |
| Labor force $l-efficiency$ workers share | $\overline{\gamma}$ | 0.38 |
| Unemployment share $l-efficiency$ | $\gamma_{ss}$ | 0.52 |
| Employment share $l-efficiency$ | $\xi_{ss}$ | 0.36 |
| Relative productivity of high-/low-efficiency workers | | 1.16 |
| Relative productivity of employed/unemployed workers | | 1.04 |
| Interview matching function efficiency | $\psi$ | 0.743 |
| Disutility of labor hours | $\ell$ | 8.5 |
| Value of home production | $w^u$ | 0.057 |
| Vacancy posting cost | $\kappa$ | 0.036 |
| Inverse of labor hours supply elasticity | $\chi$ | 1 |
| Relative risk aversion | $\sigma$ | 1 |
| Unemployment elasticity of matching function | $\alpha$ | 0.6 |
| Workers' share of surplus | $\eta$ | 0.6 |
| Steady-state separation rate | $\rho_{ss}$ | 7.45% |
| Exogenous separation rate | $\rho^x$ | 5.8% |
| Endogenous separation/screening rate | $\rho^n_{ss}$ | 4.9% |
| AR(1) parameter for exogenous shocks | $\rho_z$ | 0.7 |
| Price elasticity of retail goods demand | $\varepsilon$ | 11 |
| Discount factor | $\beta$ | 0.99 |
| Average retail price duration (quarters) | $\frac{1}{1-\omega}$ | 4 |

Table 1: Baseline Parameterization. Average productivity of high- and low-efficiency worker-hours is given by $\phi^h$ and $\phi^l \int_0^1 a^l_i dF(a^l_i)$. U.S. unemployment rate for low- and high-efficiency workers from the classification of workers based on employment spell duration from LEHD data taken from Gregory, et. al. (2021). See the Appendix for details.

Table 2: Outcome for Alternative Policies

| | Alternative policies | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Variables in rule | $\pi$ | $\pi$, $U$ | $\pi_t$, $\rho$ | $\pi$, $U^l$ | $\pi$ |
| Response coefficients. | 1.5 | $1.5, -0.39$ | $1.5, -0.015$ | $1.5, -0.5$ | |
| Target reduction | | 25% ($U$) | 25% ($U$) | 50% ($U^l$) | $\pi = 0$ |
| 1) Output loss | 32.78% | 26.14% | 27.70% | 20.03% | 27.09% |
| 2) $l$-unemployment loss | 24.56% | 17.09% | 19.18% | 10.53% | 19.01% |
| 3) $h$-unemployment loss | 3.04% | 2.48% | 2.61% | 1.97% | 2.56% |
| 4) Inequality ratio | 8.07 | 6.88 | 7.34% | 5.35 | 7.42 |
| 5) Sacrifice Ratio | 0.18 | 0.11 | 0.04 | 0.25 | 0 |

Table 2: All policies given by (17) with choice of $x_t$ specified under column number. Policies in column 2 and 3 are parameterized to reduce the rise in unemployment in the first period of the COVID shock by 25%. Policy in column 4 is parameterized to reduce by 50% the rise in $l$-workers unemployment rate in the first period of the COVID shock. Output and unemployment losses by worker-types, the unemployment inequality ratio, and the sacrifice ratio are defined in the text.