Destabilizing Search Technology*

Tristan Potter†
Drexel University

Click here for the latest version.

June 24, 2022

Abstract

Modern search technologies enable workers to monitor—and thus quickly apply to—newly posted jobs. I conceptualize search as a monitoring decision and study the implications for labor market dynamics. The central insight is that monitoring leads to a novel source of strategic complementarities in search decisions, which results in multiple equilibria that can exert a destabilizing force on the labor market. Strategic complementarities arise because workers who actively monitor new job postings are able to apply *before* those who do not. This leads to a rat race for jobs in which the belief that others are monitoring new postings necessitates doing the same in order to avoid falling to the back of the queue. I show that this mechanism leads to multiple equilibria in a stylized monitoring game and then embed the game in a quantitative macroeconomic model of the labor market. With a plausibly elastic job creation process (i.e., away from the free-entry limit), multiplicity arises in the quantitative model. The model provides (i) a theory of belief-driven fluctuations in labor supply that can permanently alter the path of the economy, (ii) a mechanism through which transitory demand shocks can permanently affect labor supply, and (iii) an account of the recovery from the Great Recession, during which a historically tight labor market coexisted with weak wage growth—observations difficult to reconcile with traditional models. I document two facts that are supportive of the model and its implications.

Keywords: Search and matching, online job search, multiple equilibria
JEL Classification: E24, J64, E71, O33

*Thanks to Jim Albrecht, Andre Kurmann, Guido Menzio, Edouard Schaal, Susan Vroman, Ronald Wolthoff, and various conference and seminar participants for helpful comments and suggestions on earlier versions of this paper.
†School of Economics, LeBow College of Business, Drexel University. Email: tristan.l.potter@drexel.edu.

1 Introduction

Since the turn of the century, online job search has become ubiquitous in the U.S. labor market. Between 2000 and 2010, the share of unemployed workers in the U.S. using the internet to search for work increased from 25% to 75% (Faberman and Kudlyak, 2016; Kuhn and Mansour, 2013). Further, as of 2013, 70% of all job openings in the U.S. were posted online (Carnevale et al., 2014). Understanding the consequences of this shift to online search requires understanding how various online search technologies affect the matching process and thus equilibrium labor market dynamics. Motivated by these questions, a growing literature has begun studying the channels through which such technologies affect the labor market. Recent examples include work considering how declining search costs can render workers more selective (Menzio and Martellini, 2020); how online recruiting can improve screening of applicants (Pries and Rogerson, 2022); and how online matching contributes to the existence of phantom vacancies (Cheron and Decreuse, 2017; Albrecht et al., 2017).[1]

This paper studies a feature of online search technologies not previously explored in the literature: Monitoring technologies that enable job seekers to find and apply to jobs soon after they are posted. A number of technologies facilitate monitoring: For instance, personal computers and broadband connections enable job seekers to easily and frequently check for new listings; job search engines sort listings by the date on which they were posted, thus facilitating finding the most recent; and matching platforms offer "job alerts" to notify job seekers of new listings that match their profiles.[2] An important consequence of such technologies is that they allow workers who actively monitor new vacancies to see and apply for these jobs *before* those who do not, a dimension of the search decision not present in traditional models. This paper conceptualizes search as a monitoring decision—that is, as a decision governing how quickly a worker is able to see and apply to newly posted vacancies. There are two main results: (i) Monitoring decisions are characterized by strategic complementarities that give rise to multiple equilibria, and (ii) the resulting multiplicity can destabilize the economy by exposing the labor market to self-fulfilling fluctuations in beliefs among jobless workers.

The first main result—that monitoring technologies can lead to multiple equilibria—can be understood in a stylized one-period game in which u unemployed workers compete for $v < u$ vacant jobs. I operationalize the notion of search as a monitoring decision by assuming that workers choose between either matching in the morning (when jobs are first posted) or enjoying a day of leisure before matching with remaining vacancies in the evening. This is a simple way to capture the fact that a decision to monitor new vacancies confers a first-mover advantage to workers who do so. Suppose, for now, that matching is frictionless in both the morning and the evening, and let w denote the value of finding a job, $\theta \equiv v/u$ the probability of finding a job if a worker applies to jobs at the same time as other workers (i.e., in either the morning by monitoring or the evening by waiting), and κ the value of leisure. Table 1 depicts this game, with a particular worker's choice represented by the rows and the (symmetric) choices of all other workers depicted by the columns. Potential equilibria are represented by diagonal entries in the matrix.

Suppose all other workers wait to search in the evening. If an individual worker does the same,

[1] See the discussion of related literature below for a more complete list of contributions to this literature.

[2] As early as 2003, Kuhn (2003) identified the potential importance of job alerts, which he refers to as "electronic job search agents." Job alerts are now offered by virtually all major job search platforms. For example, Monster.com tells searchers that they can "Get relevant jobs matching your profile and criteria straight in your inbox with our Free Job Alerts," and visitors to Indeed.com are greeted by a pop-up inviting them to enter their email address to "Be the first to see new jobs in [your location]."

Table 1: Monitoring game

Worker i	Everybody else: Monitor	Everybody else: Wait
Monitor	θw	w
Wait	κ	$\kappa + \theta w$

she enjoys a day of leisure and competes with other workers for jobs in the evening, yielding an expected payoff of $\kappa + \theta w$. On the other hand, if she decides to apply for jobs as soon as they are available in the morning, she gets a first-mover advantage and is thus assured to find a job, but forgoes leisure, giving her a payoff of w. Instead, suppose all other workers monitor and thus apply for jobs as soon as they become available in the morning. If our unemployed worker does the same, she now must compete with these workers in the morning and receives no leisure, yielding an expected payoff of θw. If instead she enjoys leisure and waits, there are no remaining jobs left in the evening, so she only receives κ.

It is straightforward to show that this game has two symmetric equilibria for $\kappa \in [(1-\theta)w, \theta w]$. The possibility of multiplicity arises because of strategic complementarities in monitoring decisions: If a worker believes that others are waiting until the evening to search, then forgoing leisure to actively monitor the arrival of jobs in the morning is unnecessary because jobs will still be available in the evening; the worker should simply wait like everyone else. By contrast, if a worker believes that others are waking up early to monitor new postings, then doing the same becomes necessary to avoid falling to the back of the queue. Furthermore, notice that the condition for multiplicity is increasingly likely to hold in a tight labor market: As $\theta \to 1$, $\kappa < w$ guarantees multiple equilibria. This is because, regardless of when others search, the value of doing the same (that is, playing the equilibrium strategy) is greater when there is greater competition for workers (i.e., when θ is high). On the other hand, the value of searching before or after everyone else (that is, deviating from equilibrium play) is not directly affected by the degree of competition for workers: Monitoring in the morning when others wait until the evening *guarantees* finding a job, whereas waiting until the evening when others search in the morning *precludes* finding a job. It follows that a tight labor market renders playing the equilibrium strategy relatively attractive, and thus sustains multiplicity.

Section 2 formalizes this monitoring game. The environment is richer than the environment embodied in Table 1, but is still static and sufficiently stylized to offer insight into several important implications of interpreting search as a monitoring decision.[3] First, monitoring leads to an aggregate match function with increasing returns. Second, increasing returns can lead to multiplicity, and the conditions under which this occurs are much like the conditions described above—a tight labor market exposes the economy to multiplicity, while a slack one precludes it. And third, no such multiplicity arises in an identical environment in which search is modeled in the traditional way—that is, as a decision of intensity rather than one of timing, as implied by monitoring.

The second main result of the paper is that the multiplicity introduced by the availability of monitoring technologies can destabilize the economy. To demonstrate this point, Section 3 embeds a version of the monitoring game described above in a quantitative equilibrium model of the labor

[3]The game in Table 1 is a limiting case of the game in Section 2.

market. As in the game above, searchers actively monitor the arrival of vacancies while other unemployed workers only have access to a standard frictional matching process that operates at the end of each period. Monitoring technologies in the quantitative model thus function to both enable workers who monitor to match quickly (as highlighted in the model and discussion above) and to reduce matching frictions. The model also features dynamic processes for unemployment and vacancies, bargained wages, and an entrepreneurial sector that gives rise to a flexible model of vacancy creation that nests various entry processes considered in the literature as special cases (e.g., free entry and "Diamond entry"). When no workers monitor the arrival of new vacancies, the model reduces to a standard search and matching model of the labor market in the Diamond-Mortensen-Pissarides (DMP) tradition,[4] augmented with a generalized vacancy-creation process.

Two important features of the quantitative model will tend to weaken the source of multiplicity that pervades the monitoring game in Table 1: endogenous wages and entry of firms.[5] In the quantitative model, as more workers choose to monitor the arrival of vacancies, wages fall. This happens because if workers believe that other workers are monitoring, they expect to have to do the same if they become unemployed—and thus to have to forgo leisure—in order to avoid falling to the back of the queue. This weakens their bargaining position and depresses wages. Depressed wages, in turn, stimulate job creation, which alleviates congestion among those at the back of the queue for jobs, and thus reduces the urgency of monitoring and mitigates the increasing returns that enabled multiplicity. Thus, if job creation is sufficiently responsive to wages, multiplicity will cease to exist. Indeed, in the limiting case of free entry, there is a unique equilibrium. By contrast, when the generalized job-creation process is not too elastic, multiplicity can arise.

Section 4 calibrates the quantitative model. The calibration is standard except for the job-creation process, which I calibrate to reflect estimates of the wage elasticity of labor demand in empirical studies. The calibrated model has three steady states, across which wages and unemployment are positively correlated. This observation reflects differences in labor supply sustained by self-confirming beliefs among unemployed workers. Specifically, each steady state is associated with different beliefs among workers about how dire their job-finding prospects will be if they do not actively monitor the arrival of vacancies and are thus relegated to the back of the queue for jobs. Different beliefs induce different monitoring decisions that, in equilibrium, confirm the beliefs that led to them. Moreover, the differences in beliefs across steady states, and the associated monitoring decisions, have normative implications: In the baseline calibration, social welfare is higher in the low-monitoring, high-unemployment steady state than in the high-monitoring, low-unemployment steady state, reflecting strong negative congestion externalities from monitoring that are not fully offset by the matching efficiency and job creation gains that monitoring brings. That this can occur is noteworthy because it reverses the Pareto-ranking in the canonical treatment of multiplicity in Diamond (1982a), in which high-activity equilibria Pareto dominate low-activity ones.

How can monitoring technologies destabilize the labor market? Section 5 considers this question. I show that, in addition to having multiple steady-state equilibria, the model also implies multiple *dynamic* equilibria, in the sense that, from a range of initial states, the economy can converge to different steady states. This can destabilize the economy in two ways. First, it implies that changes in beliefs among workers can permanently change the path of the economy, from a trajectory

[4]Diamond (1982b); Mortensen (1982); and Pissarides (1985).

[5]Recall that in the game in Table 1, both θ and w are treated as parameters, whereas both are endogenized in the quantitative model.

converging to one steady state to a trajectory converging to another. In this way, the model provides a novel theory of endogenous labor supply shocks. Second, under a simple equilibrium selection criterion, the quantitative model also implies that when the unemployment rate is sufficiently high, there is only one equilibrium path—the path leading to the high-monitoring, low-unemployment steady state—while multiple paths remain for the range of unemployment rates associated with normal business cycle fluctuations. This, in turn, implies that a sufficiently large adverse demand shock can force the economy onto a path with permanently higher monitoring than the path it began on. Put more simply, large transitory demand shocks can permanently affect labor supply.

This second observation can potentially offer insight into the recovery from the Great Recession. After unemployment peaked at 10% in late 2009, it ultimately fell substantially *below* its pre-recession level, while the vacancy rate rose substantially *above* its pre-recession level. Meanwhile, as was noted at the time,[6] even as late as 2019, wage growth remained weak relative to what would have been expected based on vacancy and unemployment data. These observations are difficult to square with both traditional models of the labor market and models of multiple steady states generated by demand-side mechanisms. In the present model, a demand shock that drives the economy to 10% unemployment forces the economy to the high-monitoring steady state. This implies not only a slow recovery, but also one that overshoots the original levels of unemployment and vacancies from which the economy started, and in which wages never fully recover. Notwithstanding permanently lower unemployment, because of the high level of monitoring, the steady state towards which the economy converges is Pareto-dominated by the low-monitoring steady state from which it began.

In Section 6, I document two features of the data that are supportive of monitoring and its implications. First, I use high-frequency panel data on unemployed workers' search decisions to document that workers who use the internet for search are significantly more likely to search on any given day than those who do not, even after controlling for time spent on job search and job-seeker fixed effects—that is, online search is less "lumpy" than traditional search. This is consistent with online job search allowing workers to monitor the arrival of new job openings and thus apply for new jobs *as they arrive*, rather than devoting, e.g., a single day of the week to job search. Second, I revisit the macroeconomic SVAR literature that seeks to identify labor supply shocks in aggregate data. Using a theoretically motivated sign restriction approach to identification, I show that the contribution of labor supply shocks to aggregate fluctuations at all horizons increased markedly around the turn of the century. This is consistent with the emergence of endogenous belief-driven fluctuations in job search driven by growth in monitoring technologies beginning in the early 2000s. While these pieces of evidence are not conclusive, they are, respectively, supportive of monitoring and its implications.

Related literature. This paper lies at the intersection of the macroeconomic literature that examines the possibility of multiple equilibria and self-fulfilling fluctuations arising in the labor market, the growing literature on how the internet affects job search and the matching process, and the macroeconomic literature that considers the role of labor supply shocks in aggregate fluctuations.

The paper is most closely related to the macroeconomic literature interested in mechanisms through which multiplicity can arise in models of the labor market. The seminal contribution in this literature is Diamond (1982a), with important subsequent work including Howitt and McAfee (1992); Pissarides (1992); Schaal and Taschereau-Dumouchel (2016); Kaplan and Menzio (2016);

[6] For example, see John Robertson's discussion in a post on the Atlanta Fed's Macroblog on November 25, 2019.

Sterk (2016); Eeckhout and Lindenlaub (2018); and Acharya et al. (2018). While the work since Diamond (1982a) considers various mechanisms, much of this work focuses on the role of firms' hiring or employment decisions in generating multiplicity. In this paper, strategic complementarities arise from the search decisions of unemployed workers rather than from firms' hiring or employment decisions.[7] The fact that multiplicity arises through a labor supply channel allows my model to account for several recent empirical phenomena that are otherwise difficult to explain.[8]

The paper also relates to a large and growing literature regarding how the internet affects matching and the labor market. In addition to the papers mentioned previously, contributions to this literature include Autor (2001); Freeman (2002); Kuhn (2003); Kuhn and Skuterud (2004); Stevenson (2009); Kuhn and Mansour (2013); Kroft and Pope (2014); Faberman and Kudlyak (2016); Faberman and Kudlyak (2019); Belot et al. (2019); and Davis and Samaniego de la Parra (2020). Two of these papers are of particular relevance in the present context. First, Kuhn (2003) is among the first to explicitly identify what he refers to as "electronic job search agents"—that is, a job board feature that "emails a worker when a new vacancy satisfying worker-supplied criteria is posted." As discussed above, this feature of job boards is essential for enabling workers to monitor vacancies. Second, Davis and Samaniego de la Parra (2020) provide direct empirical evidence of "application bunching" in online matching platforms—an empirical phenomenon in which new vacancies receive a large number of applications immediately after posting. This is precisely what is anticipated by the model of monitoring developed in this paper. Also related are models of stock-flow matching (Coles and Smith, 1998). In the present paper, the availability of monitoring technologies affords workers a choice between having immediate access to the inflow of vacancies versus waiting and only being able to apply for the leftover stock of jobs.

Finally, the paper relates to a macroeconomic literature that begins with Shapiro and Watson's (1988) evidence on the role of labor supply shocks in macroeconomic fluctuations. Important contributions to this literature include Blanchard and Diamond (1989); Chang and Schorfheide (2003); Peersman and Straub (2009); and more recently Foroni et al. (2018), all of whom find an important role for labor supply shocks—in many cases, even at business cycle frequencies. This paper provides a novel, micro-founded theory of endogenous labor supply shocks. Furthermore, I build on the finding of Foroni et al. (2018) that labor supply shocks have become *increasingly* important in recent decades by providing evidence that the contribution of labor supply shocks at all frequencies is substantially larger in the post-2000 sample than in the pre-2000 sample. This evidence is consistent with a mechanism rooted in the existence of online search technology.

Outline of paper. The paper proceeds as follows. Section 2 studies a stylized monitoring game to illustrate the mechanism. Section 3 embeds the game in a quantitative model of the labor market. Section 4 calibrates the model and considers its steady states. Section 5 turns to dynamics, focusing on how belief-driven fluctuations can emerge, and the model's implications for the Great Recession. Section 6 documents two pieces of evidence consistent with the model and its implications.

[7] In Eeckhout and Lindenlaub (2018), multiplicity arises due to strategic complementarities between firms' vacancy posting decisions and workers' on-the-job search decisions. Thus, their model also highlights an important role for job search, although the mechanism is fundamentally different from the one I study.

[8] Following Diamond (1982a), an empirical literature surveyed by Petrongolo and Pissarides (2001) has generally failed to find robust evidence of increasing returns in aggregate matching functions. However, because much of this literature preceded the widespread adoption of monitoring technologies, these findings do not bear on whether such technologies might function as a *modern* source of increasing returns.

2 Monitoring Game

The central insight of this paper is that monitoring technologies can lead to strategic complementarities in search decisions and thus multiple equilibria. To illustrate the mechanism, I begin by considering a stylized game among unemployed workers with access to a monitoring technology. The analysis yields three main insights. First, understanding search as monitoring leads to an aggregate matching function with increasing returns. Second, increasing returns can give rise to multiple equilibria, and are particularly likely to do so in a tight labor market. And third, multiplicity cannot arise in an identical environment with a more traditional model of search.

2.1 Search as monitoring

The technologies discussed in the Introduction—new job alerts, job search engines that filter results by date of posting, and even personal computers and broadband connections—can all be understood as technologies that function to reduce the *fixed costs* of finding new job openings for which to apply. For example, prior to the existence of such technologies, finding new job openings might have necessitated driving to a job center 30 minutes away—a job center that would eventually migrate online and possibly even actively notify workers of newly listed jobs. High fixed costs, in turn, make frequent search costly: Clearly it would not be optimal for a worker to drive to the job center to check for new jobs several times each day; more likely, they will drive to the job center, say, once per month. Put differently, for a given planning horizon, sufficiently high fixed costs will tend to render the decision of how frequently to search trivial, leading to a corner solution in which search decisions are infrequent and appear "lumpy."[9]

This intuition suggests that the advent of monitoring technologies that reduce the fixed costs of search will lead to a non-trivial decision regarding how frequently to check for newly posted jobs.[10] This is the decision that I study. To operationalize the decision of search frequency in a tractable way, throughout the paper I assume that workers make a discrete choice between arbitrarily frequent search (that is, monitoring) and simply checking for jobs at the end of each period. An appealing feature of this approach is that, if vacancies arrive sequentially and at random times throughout each period (and are able to make offers relatively quickly), it implies a frictionless matching process between new vacancies and the workers who choose to monitor those vacancies, and thus provides an explicit foundation for the common claim that online search reduces matching frictions. This feature of the model has important implications for welfare to which I return in Section 4.

2.2 Environment

Consider a one-period game among u unemployed workers competing to match with $v < u$ vacancies that are posted sequentially at random times throughout the period. Workers who are matched by the end of the period receive payoff w and workers who are unmatched receive payoff $b < w$.

[9]This does not imply that there are not other relevant dimensions of the search decision in the presence of high fixed costs of finding new openings, such as how many of those jobs to apply to or how much "effort" to put into an application. High fixed costs thus bear on the number of distinct episodes of search rather than the intensity of search within each episode. To keep the analysis tractable, I focus on the monitoring component of the search decision.

[10]In Section 6 I document that search is significantly less lumpy for job seekers who report using the internet for search. This is consistent with monitoring technologies reducing fixed costs.

2.2.1 Monitoring decision

At the start of the period, each worker i draws flow utility of leisure (or home production) κ_i from distribution F with support $[\underline{\kappa}, \bar{\kappa}]$. After observing κ_i, workers decide how to spend their time: They can either actively monitor the arrival of vacancies throughout the period, thus gaining a first-mover advantage in submitting applications, or enjoy leisure/home production and wait until the end of the period to try to find work. Workers choosing to monitor match with probability p^m but forgo leisure/home production, while workers choosing to wait match with probability $p^w \leq p^m$ but enjoy flow utility κ_i. The values of monitoring and waiting for worker i, respectively, are thus

$$
\begin{aligned}
U^m \quad &= p^m w + (1 - p^m) b & (1)\\
U^w(\kappa_i) &= p^w w + (1 - p^w) b + \kappa_i. & (2)
\end{aligned}
$$

Workers monitor if and only if the value of monitoring exceeds the value of waiting: $U^m > U^w(\kappa_i)$. Equations (1) and (2) reveal that monitoring decisions are characterized by a cutoff rule such that a worker drawing $\kappa_i < \kappa_i^*$ will choose to monitor while a worker drawing $\kappa_i \geq \kappa_i^*$ will choose to wait, where (when interior) κ_i^* is defined by $\kappa_i^* \equiv \{\kappa_i | U^m = U^w(\kappa_i)\}$. Using (1) and (2),

$$
\kappa_i^* = \begin{cases} \underline{\kappa} & \text{if } \big(p^m - p^w\big)(w - b) < \underline{\kappa} \\ \big(p^m - p^w\big)(w - b) & \text{if } \big(p^m - p^w\big)(w - b) \in [\underline{\kappa}, \bar{\kappa}] \\ \bar{\kappa} & \text{if } \big(p^m - p^w\big)(w - b) > \bar{\kappa} \end{cases} \qquad (3)
$$

where, implicitly, the match rates p^m and p^w depend on the equilibrium behavior of other workers (and the number of vacancies) in the economy. I next derive expressions for these two match rates.

2.2.2 Matching

The preceding description implies that matching effectively occurs in two phases: A monitoring phase (in which monitoring workers match with arriving vacancies) and an aftermarket (in which all vacancies and workers that did not match in the monitoring phase are matched). Because vacancies arrive sequentially throughout the period and monitoring allows workers to immediately observe and apply for new jobs, matching in the monitoring phase is frictionless. Letting κ^* denote the (symmetric) equilibrium monitoring cutoff, $F(\kappa^*)u$ is the number of monitoring workers, and the number of matches in the monitoring phase is

$$
\mathring{m} = \min\big\{v, F(\kappa^*)u\big\} \qquad (4)
$$

implying that the match rate for workers in the monitoring phase is $\mathring{p} \equiv \frac{\mathring{m}}{F(\kappa^*)u}$.

For simplicity, I assume that the aftermarket is likewise frictionless.[11] Thus, the number of matches in the aftermarket is given by

$$
\begin{aligned}
\hat{m} &= \min\{v - \mathring{m}, u - \mathring{m}\} & (5)\\
&= v - \mathring{m} & (6)
\end{aligned}
$$

[11] The frictionless property of the monitoring phase is a consequence of monitoring assuming vacancies are posted randomly throughout the period and firms can quickly make offers, whereas the frictionless property of the aftermarket is an assumption that I relax in Section 3.

implying that the match rate for workers in the aftermarket is $\hat{p} \equiv \frac{\hat{m}}{u-\mathring{m}}$.[12]

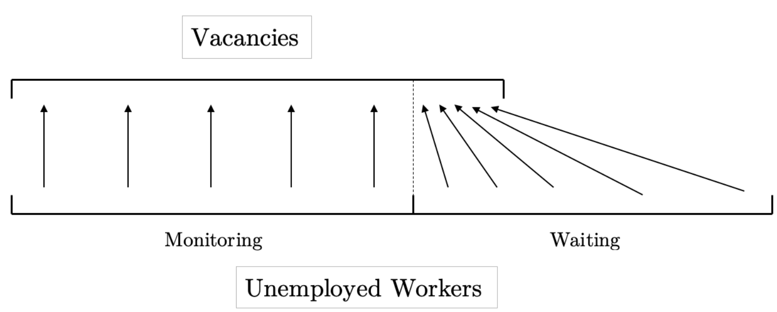
Workers who choose to monitor are able to match during the monitoring phase and, failing that, in the aftermarket.[13] Workers who choose to wait are only able to match in the aftermarket. This implies that the match rates for monitoring workers and waiting workers that appear in (3) are, respectively, given by $p^m = \mathring{p} + (1-\mathring{p})\hat{p}$ and $p^w = \hat{p}$. Using these together with (4) and (6), and defining market tightness as $\theta \equiv v/u < 1$, it is straightforward to show that the match rates for monitoring workers and waiting workers, respectively, can be written as

$$p^m = \begin{cases} 1 & \text{if } \; F(\kappa^*) \leq \theta \\ \frac{\theta}{F(\kappa^*)} & \text{if } \; F(\kappa^*) > \theta \end{cases} \tag{7}$$

$$p^w = \begin{cases} \frac{\theta - F(\kappa^*)}{1-F(\kappa^*)} & \text{if } \; F(\kappa^*) \leq \theta \\ 0 & \text{if } \; F(\kappa^*) > \theta. \end{cases} \tag{8}$$

The model is fully described by equations (3), (7), and (8). Figure 1 depicts the matching process.

Figure 1: Matching with monitoring



Notes: The figure depicts the matching process in the monitoring game.

2.2.3 Increasing returns

Before characterizing the Nash equilibria, it is useful to reflect on the source of increasing returns. Observe that the match rates in (7) and (8) imply that the return to monitoring is

$$p^m - p^w = \begin{cases} \frac{1-\theta}{1-F(\kappa^*)} & \text{if } \; F(\kappa^*) \leq \theta \\ \frac{\theta}{F(\kappa^*)} & \text{if } \; F(\kappa^*) > \theta. \end{cases} \tag{9}$$

Equation (9) reveals that, when the labor market is tight ($F(\kappa^*) \leq \theta$), the model exhibits increasing returns in the sense that the individual benefit from monitoring relative to waiting, $p^m - p^w$, is an increasing function of the equilibrium monitoring cutoff, κ^*. As has been well known since the seminal work of Diamond (1982a), this property can expose the economy to multiple equilibria.

[12]The fact that $\min\{v - \mathring{m}, u - \mathring{m}\} = v - \mathring{m}$ follows from the maintained assumption of $v < u$.

[13]This assumption is inessential.

Monitoring technologies can lead to increasing returns and thus multiplicity because of the first-mover advantage that they confer to searchers. In a tight labor market ($F(\kappa^*) \leq \theta$), the existence of a first-mover advantage *reduces* congestion among workers who monitor ($p^m = 1$) at the expense of *increased* congestion among those who wait, all of whom must compete for a diminished stock of remaining jobs. As a consequence, as more workers monitor, the job-finding prospects of workers waiting until the aftermarket deteriorate ($\frac{\partial p^w}{\partial \kappa^*} < 0$), and can do so increasing rapidly ($\frac{\partial^2 p^w}{\partial \kappa^{*2}} < 0$) depending on the shape of F.[14] This, in turn, causes the return to monitoring to rise ($\frac{\partial (p^m - p^w)}{\partial \kappa^*} > 0$), and to potentially do so increasingly rapidly ($\frac{\partial^2 (p^m - p^w)}{\partial \kappa^{*2}} > 0$). This dynamic is the fundamental reason why monitoring technologies can give rise to multiple equilibria.

2.3 Nash equilibria

Inspection of (7) and (8) reveals that the match rates p^m and p^w depend on the cutoff chosen by other workers in the economy: $p^m = p^m(\kappa^*)$ and $p^w = p^w(\kappa^*)$. Accordingly, the optimal cutoff of worker i in equation (3) can be written as a best-response function: $\kappa_i^* = \kappa_i^*(\kappa^*)$. The (symmetric) Nash equilibria of the game are then the fixed points of this best-response function, i.e. the points at which $\kappa_i^*(\kappa^*) = \kappa^*$.

To characterize the set of equilibria, substitute (7) and (8) into (3) to obtain the best-response function expressed explicitly as a function of κ^*:

$$\kappa_i^*(\kappa^*) = \begin{cases} \kappa^L(\kappa^*) & \text{if} \quad F(\kappa^*) \leq \theta \\ \kappa^U(\kappa^*) & \text{if} \quad F(\kappa^*) > \theta \end{cases} \tag{10}$$

where

$$\kappa^L(\kappa^*) \equiv \begin{cases} \underline{\kappa} & \text{if} \quad \frac{1-\theta}{1-F(\kappa^*)}(w-b) < \underline{\kappa} \\ \frac{1-\theta}{1-F(\kappa^*)}(w-b) & \text{if} \quad \frac{1-\theta}{1-F(\kappa^*)}(w-b) \in [\underline{\kappa}, \bar{\kappa}] \\ \bar{\kappa} & \text{if} \quad \frac{1-\theta}{1-F(\kappa^*)}(w-b) > \bar{\kappa} \end{cases} \tag{11}$$

$$\kappa^U(\kappa^*) \equiv \begin{cases} \underline{\kappa} & \text{if} \quad \frac{\theta}{F(\kappa^*)}(w-b) < \underline{\kappa} \\ \frac{\theta}{F(\kappa^*)}(w-b) & \text{if} \quad \frac{\theta}{F(\kappa^*)}(w-b) \in [\underline{\kappa}, \bar{\kappa}] \\ \bar{\kappa} & \text{if} \quad \frac{\theta}{F(\kappa^*)}(w-b) > \bar{\kappa}. \end{cases} \tag{12}$$

Equations (10), (11) and (12) indicate that $\kappa_i^*(\kappa^*)$ is continuous, has a kink at $\kappa^c \equiv F^{-1}(\theta)$, is weakly increasing for $\kappa^* \leq \kappa^c$, and is weakly decreasing for $\kappa^* > \kappa^c$. In general, the number of equilibria will depend on the shape of the distribution function F. Nevertheless, Proposition 1 provides sufficient conditions for multiplicity (and for uniqueness).[15]

Proposition 1 (Sufficient conditions for multiplicity and uniqueness). *Suppose that $\bar{\kappa} \in (0, w-b)$ (i.e., the type who derives the most utility from leisure will monitor if guaranteed to match). Then:*

1. *There is a value $\bar{\theta} < 1$ such that for any $\theta \in (\bar{\theta}, 1)$ there are multiple equilibria.*
2. *There is a value $\underline{\theta} > 0$ such that for any $\theta \in (0, \underline{\theta})$ an equilibrium exists, it is unique, and it satisfies $\kappa^* > \kappa^c$. That is, the unique equilibrium entails high monitoring.*

[14] This will be the case, for example, if $f' \geq 0$.

[15] See Appendix A for discussion of a special case that admits a closed-form solution for the set of Nash equilibria.

Proof. See Appendix A. □

Proposition 1 establishes two important features of monitoring. First, monitoring can generate multiple equilibria. This is consistent with the discussion of increasing returns above. Second, the emergence of multiplicity is intimately related to market tightness. Precisely, multiplicity is guaranteed by a sufficiently tight labor market and precluded by a sufficiently slack one. The fundamental reason is that monitoring temporally segments the matching process, as depicted in Figure 1. As a consequence, deviations from equilibrium play entail payoffs that are independent of market tightness, whereas playing the equilibrium strategy is encouraged by a tight labor market and discouraged by a slack one, an observation that may be seen clearly in the game in Table 1.[16] This observation is the key feature of monitoring that allows for multiplicity to emerge, and one that will have important implications in the analysis of the quantitative model in Section 5.

2.4 Traditional model

The traditional way of modeling search decisions is to assume that non-searchers are simply ineffective or low-effort searchers, matching at a rate proportional to that of searchers.[17] The environment described in Section 2.2 can be adapted to consider such a model by simply replacing p^m with p^h ("high" effort) and p^w with p^l ("low" effort) and deriving the corresponding match rates. Let $z < 1$ denote effort of non-searchers relative to searchers, and continue to denote by $F(\kappa^*)$ the fraction of workers choosing to search in equilibrium. Then, the effective measure of searchers is[18]

$$\begin{aligned} s(\kappa^*) &= F(\kappa^*)u + (1 - F(\kappa^*))zu && (13) \\ &= \tilde{s}(\kappa^*)u && (14) \end{aligned}$$

where $\tilde{s}(\kappa^*) \equiv F(\kappa^*) + (1 - F(\kappa^*))z$ is average effective search effort. Maintaining the assumption of frictionless matching, the number of matches is $m = \min\{v, \tilde{s}(\kappa^*)u\}$, which implies match rates

$$p^h = \begin{cases} 1 & \text{if } \tilde{s}(\kappa^*) \leq \theta \\ \frac{\theta}{\tilde{s}(\kappa^*)} & \text{if } \tilde{s}(\kappa^*) > \theta \end{cases} \tag{15}$$

$$p^l = \begin{cases} z & \text{if } \tilde{s}(\kappa^*) \leq \theta \\ z\frac{\theta}{\tilde{s}(\kappa^*)} & \text{if } \tilde{s}(\kappa^*) > \theta. \end{cases} \tag{16}$$

Inspection of (15) and (16) in light of (13) and $z < 1$ reveals that the return to search, $p^h - p^l$, is a decreasing function of the search cutoff, κ^*, which implies that there cannot be multiplicity. This observation is formalized in Proposition 2:

Proposition 2 (No multiplicity). *There is at most one equilibrium in the traditional model.*

Proof. See Appendix A. □

Thus, understanding search as a monitoring decision is the essential feature of the environment that gives rise to multiple equilibria. This result will have a counterpart in the analysis of the quantitative model in Section 3, to which I now turn.

[16] The game in Table 1 is a special case of the model described above when $b = 0$ and F is degenerate.

[17] For example, Krusell et al. (2017) and Cairo et al. (2021) consider such models.

[18] See Appendix B for further details and discussion of the traditional model.

3 Dynamic Model

To study the implications of monitoring for equilibrium labor market dynamics, I embed a variation on the static game from Section 2 within a quantitative macroeconomic model of the labor market. The model is fully dynamic and features a generalized version of the matching process in Section 2, bargained wages, and a flexible model of vacancy creation that nests various special cases considered in the literature. In the absence of monitoring, the model reduces to a standard DMP-style model of the labor market with a generalized job-creation process.

3.1 Environment

Time is discrete and runs forever. The economy is populated by a unit measure of ex ante identical workers who are either employed or unemployed and a fixed measure of entrepreneurs who create vacancies. Workers and entrepreneurs all discount the future with discount factor β, have perfect foresight with respect to the evolution of aggregate variables,[19] and seek to maximize the present discounted value of lifetime utility.

Important elements of the model below, namely those relating to workers' decisions and the matching process, are similar to the simple game described previously in Section 2. Nevertheless, in order for the analysis in this section to be fully self-contained, I provide a complete description of these elements of the model below.

3.1.1 Accounting

Let u_t denote the total number of unemployed workers in a period and v_t the total number of vacancies in a period.[20] Unemployed workers either actively monitor the arrival of vacancies throughout the period (there are u_t^m such workers) or wait until the end of the period to match (there are u_t^w such workers). Vacancies are either newly posted by an entrepreneur (there are v_t^n such vacancies) or were posted in a previous period and are therefore old (there are v_t^o such vacancies). This implies the following identities:

$$u_t = u_t^m + u_t^w \tag{17}$$

$$v_t = v_t^n + v_t^o. \tag{18}$$

Unemployed workers and vacancies are matched through a frictional matching process that I describe in further detail below. Matches become productive in the period after they are formed and are destroyed with probability δ at the end of each period (including matches that have formed but have not yet become productive). Letting m_t denote the total number of matches in a period, unemployment and vacancies evolve according to the following laws of motion:

$$u_{t+1} = u_t + \delta(1 - u_t) - (1 - \delta)m_t \tag{19}$$

$$v_{t+1} = (1 - \delta)(v_t - m_t) + v_{t+1}^n. \tag{20}$$

[19] In Section 5, I consider the effects of unanticipated "MIT shocks" to the economy.

[20] While I use the term "unemployment" throughout the paper, the model should be understood as referring more broadly to any non-employed workers who have some attachment to the labor market.

Equations (19) and (20) are standard and describe the aggregate dynamics of unemployment and vacancies in the model.

3.1.2 Matching

As in Section 2, unemployed workers choosing to monitor new job postings are able to observe (and apply for) new jobs as soon as they become available—and thus before other unemployed workers who are *not* actively monitoring new postings. Monitoring thus results in a temporally segmented matching process that can be thought of as taking place in two phases: A monitoring phase (which lasts for the duration of the period) followed by an aftermarket phase (at the end of the period).

Monitoring. Newly created vacancies arrive randomly throughout each period. During the monitoring phase, as newly posted vacancies arrive to the market, they are matched with workers who have chosen to monitor those arrivals. Because vacancies arrive sequentially throughout the period and monitoring allows workers to see new vacancies immediately, the matching process in the monitoring phase is frictionless and the total number of matches is given by[21]

$$\mathring{m}_t = \min\{v_t^n, u_t^m\}. \tag{21}$$

The corresponding match rates for workers and vacancies in the monitoring phase are, respectively,

$$\mathring{p}_t = \mathring{m}_t / u_t^m \tag{22}$$
$$\mathring{q}_t = \mathring{m}_t / v_t^n. \tag{23}$$

Aftermarket. At the end of each period, after all new vacancies have been posted, the aftermarket opens. In the aftermarket, all remaining unmatched workers (monitoring workers who failed to match while monitoring and workers who did not monitor) match with all unmatched vacancies (newly posted vacancies that failed to match with monitoring workers and old vacancies that were posted in previous periods and thus were not observed by monitoring workers). Thus, the total number of matches in the aftermarket is given by[22]

$$\hat{m}_t = \mu\big(v_t - \mathring{m}_t, u_t - \mathring{m}_t\big). \tag{24}$$

where μ is a matching function. The corresponding match rates for workers and vacancies in the aftermarket are, respectively,

$$\hat{p}_t = \hat{m}_t / (u_t - \mathring{m}_t) \tag{25}$$
$$\hat{q}_t = \hat{m}_t / (v_t - \mathring{m}_t). \tag{26}$$

The preceding implies that the total number of matches in the economy is given by

$$m_t = \hat{m}_t + \mathring{m}_t \tag{27}$$

with corresponding average match rates $p_t = m_t/u_t$ and $q_t = m_t/v_t$. Notice that the matching process described above reduces to a standard matching model in the absence of monitoring tech-

[21]I denote variables corresponding to the monitoring phase with circles, e.g. $\mathring{x}$.

[22]I denote variables corresponding to the aftermarket with hats, e.g. $\hat{x}$.

nology: That is, if $\mathring{m}_t = 0$, then the total number of matches is just $m_t = \mu(v_t, u_t)$. Thus, the model is simply a generalization of a traditional matching model with no search decision, in which monitoring technology affords searchers a first-mover advantage. I next consider workers' optimal choice to use this monitoring technology.

3.1.3 Workers

Employment and unemployment. Workers are either employed or unemployed. Employed workers receive wage w_t at the beginning of each period. Unemployed workers receive income $b_t = bw_t$ at the beginning of each period and decide whether or not to monitor new job postings.

Workers who do not monitor new postings draw i.i.d. flow utility κ_{it} (e.g., leisure or home production) from distribution F and are only able to match in the aftermarket at the end of the period. Workers who monitor new postings, on the other hand, do not get to draw flow utility κ_{it}, but have the opportunity to match during the monitoring phase with newly posted vacancies in addition to being able to match in the aftermarket if they fail to match while monitoring. Thus, once again letting p^w_t denote the match rate for workers who wait until the aftermarket rather than monitoring and p^m_t the match rate for workers who monitor, the preceding implies

$$p^m_t = \mathring{p}_t + (1 - \mathring{p}_t)\hat{p}_t \tag{28}$$

$$p^w_t = \hat{p}_t. \tag{29}$$

Monitoring decision. Let W_t denote the value of entering period t employed, U_{it} the value of entering period t unemployed with draw κ_{it}, U^m_t the value of being unemployed and monitoring new postings throughout the period, and U^w_{it} the value of being unemployed and waiting until the aftermarket to match. Then the description of the environment above implies

$$W_t \ = w_t + \Big[\delta\beta\mathbb{E}U_{it+1} + (1-\delta)\beta W_{t+1}\Big] \tag{30}$$

$$U_{it} \ = b_t + \max\Big\{U^m_t, U^w_{it}\Big\} \tag{31}$$

$$U^m_t = p^m_t\Big[\delta\beta\mathbb{E}U_{it+1} + (1-\delta)\beta W_{t+1}\Big] + (1-p^m_t)\Big[\beta\mathbb{E}U_{it+1}\Big] \tag{32}$$

$$U^w_{it} \ = p^w_t\Big[\delta\beta\mathbb{E}U_{it+1} + (1-\delta)\beta W_{t+1}\Big] + (1-p^w_t)\Big[\beta\mathbb{E}U_{it+1}\Big] + \kappa_{it} \tag{33}$$

where the expectation operator reflects uncertainty regarding future (i.i.d.) draws of κ_{it}. As may be seen in (32) and (33), an unemployed worker's decision about whether or not to monitor the arrival of new postings entails a tradeoff between a higher probability of matching if she chooses to monitor ($p^m_t \geq p^w_t$), and higher flow utility if she chooses not to monitor ($b_t + \kappa_{it} \geq b_t$).

Unemployed workers choose to monitor the arrival of new postings if and only if the value of doing so exceeds the value of not doing so, i.e. $U^m_t > U^w_{it}$. Because κ_{it} is i.i.d., (32) and (33) imply that the monitoring decision takes the form of a cutoff rule such that workers choose to monitor if and only if the value of leisure/home production exceeds a threshold defined by the value of κ_{it} such that $U^m_t = U^w_{it}$. Using (30)-(33), it is straightforward to show that this cutoff, κ^*_{it}, is given by

$$\kappa^*_{it} = \beta(1-\delta)(p^m_t - p^w_t)\mathbb{E}\Big[W_{t+1} - U_{it+1}\Big]. \tag{34}$$

In a symmetric equilibrium $\kappa_{it}^* = \kappa_t^*$ for all i. Thus, the number of monitoring and waiting workers in period t are given, respectively, by

$$u_t^m = F(\kappa_t^*)u_t \tag{35}$$
$$u_t^w = (1 - F(\kappa_t^*))u_t. \tag{36}$$

3.1.4 Entrepreneurs and job creation

I consider a generalized job-creation process that parameterizes the elasticity of job creation and thus nests free entry and inelastic job-creation processes as limiting cases.[23] Specifically, there is a fixed number η of entrepreneurs in the economy, each of whom can invest in the creation of a new job (i.e., create a vacancy). During each period t, each such entrepreneur j receives an opportunity to draw i.i.d. sunk investment cost ξ_{jt} from distribution G. If an entrepreneur decides to invest, it pays the sunk cost and creates a new vacancy with value V_t^n. Entrepreneurs undertake the investment if the expected value of creating a new vacancy exceeds the sunk investment cost, i.e., if $V_t^n > \xi_{jt}$. Thus, the number of newly created vacancies in the economy, v_t^n, is determined by

$$v_t^n = \eta G\big(V_t^n\big). \tag{37}$$

Letting q_t^n denote the match rate for new vacancies and q_t^o the match rate for old vacancies, the value of a new vacancy, an old vacancy, and a filled job, respectively, are given by

$$V_t^n = -c + \beta(1-\delta)\Big[q_t^n J_{t+1} + (1 - q_t^n)V_{t+1}^o\Big] \tag{38}$$
$$V_t^o = -c + \beta(1-\delta)\Big[q_t^o J_{t+1} + (1 - q_t^o)V_{t+1}^o\Big] \tag{39}$$
$$J_t = y_t - w_t - \tau_t + \beta(1-\delta)\Big[J_{t+1}\Big] \tag{40}$$

where y_t is match output, τ_t is a lump-sum tax used to finance UI payments b_t to unemployed workers, and c is the cost of maintaining a vacancy. Note that the description of the matching process above implies that q_t^n and q_t^o are given by

$$q_t^n = \mathring{q}_t + (1 - \mathring{q}_t)\hat{q}_t \tag{41}$$
$$q_t^o = \hat{q}_t. \tag{42}$$

Considering a generalized job-creation process, such as that described in (37), is important for the existence of multiple equilibria. In particular, in the free-entry limit, vacancy creation is highly responsive to workers' monitoring decisions, and because the incentive to monitor diminishes with the number of firms in the economy (all else equal), a highly elastic job creation process precludes multiplicity. In Section 4, I calibrate the elasticity of the entry process to be consistent with empirical estimates of the wage elasticity of labor demand in the literature and also explore the implications of more elastic entry processes.[24]

[23] See Fonseca et al. (2000), Fujita and Ramey (2005), Beaudry et al. (2018), and Coles and Moghaddasi Kelishomi (2018) for examples of similar processes.

[24] If the number of entrepreneurs receiving investment opportunities in a period becomes small as the length of the period tends to zero, then the number of workers who are able to monitor vacancies in a period must also become small in order for the nature of the matching process to remain in tact in the continuous-time limit. This would

3.1.5 Wages

Wages are determined by Nash bargaining. Letting χ denote workers' bargaining power, the Nash bargained wage solves

$$(1-\chi)\Big[W_t - \mathbb{E}U_{it}\Big] = \chi\Big[J_t - V_t^o\Big]. \tag{43}$$

Equation (43) reflects the fact that, if bargaining breaks down, a vacancy is no longer new and thus the outside option for any firm is V_t^o. This implies that the wage is the same in all matches.

3.1.6 Taxes and transfers

The government is assumed to run a balanced budget, so that total transfers to unemployed workers are paid for by a lump-sum tax on firms τ_t:

$$\tau_t = \frac{bw_t u_t}{1-u_t}. \tag{44}$$

3.2 Equilibrium

Definition 1 defines a (symmetric) perfect foresight equilibrium of the dynamic model described above.

Definition 1. *A (symmetric) perfect foresight equilibrium is a sequence* $\{u_t, v_t, \kappa_{it}^*, v_t^n\}$ *such that* $\kappa_{it}^* = \kappa_t^*$ *for all* i*, and satisfying* (19),(20), (34)*, and* (37) *for all* $t \geq 1$*, given:*

1. *The definitions and equilibrium conditions in* (17)-(44)*;*
2. *Initial conditions:* $\{u_0, v_0\}$*;*
3. *Transversality conditions:* $\lim_{t\to\infty} v_t^n < \infty$ *and* $\lim_{t\to\infty} \kappa_t^* < \infty$*.*

4 Calibration and Steady States

With the complete dynamic model in hand, I turn to calibrating the model's structural parameters. I first describe the calibration of the model, then show that the calibration leads to multiple steady states, and finally show how the emergence of multiplicity depends critically on both the interpretation of search as monitoring (as opposed to a more traditional interpretation) and the elasticity of the job-creation process. At the end of the section, I discuss how social welfare varies across steady states and show that the model reverses the Pareto ranking of Diamond (1982a).

4.1 Calibration

I calibrate the model in two steps. In the first step, I consider the model with no monitoring (i.e., $\mathring{m} = 0$), and calibrate all of the standard parameters of this model (that is, all of the parameters that are relevant in a standard DMP-style model without monitoring) to either match values commonly used in the literature or long-run averages of data from prior to 2000.[25] This pins down

result, for example, from unemployed workers receiving opportunities to search periodically in the same way that entrepreneurs periodically receive opportunities to invest.

[25] The model without monitoring used for the first step of calibration is obtained by replacing the first-order condition for κ_t^* in (34) with $\kappa_t^* = 0$.

all of the parameters except for the standard deviation of the leisure/home production distribution. In the second step, I consider the model with monitoring (i.e., $\mathring{m} > 0$), and calibrate this final parameter to match a back-of-envelope estimate of the prevalence of monitoring among jobless workers.

The rationale for this two-step approach is twofold: First, because there is not a discrete point in time at which monitoring technologies became widespread—presumably they began to become increasingly prevalent around the turn of the century, but it is likely that their use continued to grow for a number of years thereafter—calibrating structural parameters to a version of the model with monitoring would entail an inherent ambiguity about which years should be targeted when choosing moments in the data (as well as simply a limited number of available years). The impact of this ambiguity can be minimized by calibrating all parameters that are unrelated to monitoring to the period before such technologies were available (i.e., pre-2000). Second, and relatedly, in order to isolate the effects of monitoring, it is useful to consider a model and calibration that is entirely standard in the absence of monitoring. The approach I take addresses both of these issues.

Below, I first describe functional forms and then describe how I calibrate individual parameters. See Appendix C for complete details.

4.1.1 Functional forms

Following Beaudry et al. (2018) and Coles and Moghaddasi Kelishomi (2018), I adopt a flexible form for the entrepreneurial investment cost function,

$$G(\xi_{jt}) = \left(\xi_{jt}/\bar{\xi}\right)^{\nu}. \tag{45}$$

The parameter ν governs the elasticity of job creation, allowing (45) to nest both free-entry and an inelastic vacancy creation process as special cases.[26] I assume that workers and vacancies that remain unmatched at the end of each period are matched via an urn-ball matching function,

$$\mu\big(v_t - \mathring{m}_t, u_t - \mathring{m}_t\big) = (v_t - \mathring{m}_t)\left[1 - e^{-\psi\frac{u_t - \mathring{m}_t}{v_t - \mathring{m}_t}}\right]. \tag{46}$$

This functional form has the virtue of remaining bounded between 0 and 1, which is particularly important when there are potentially multiple equilibria, and furthermore naturally nests the fully frictionless matching process described in Section 2 as the matching efficiency parameter ψ becomes large, a fact that I will exploit in the numerical analysis in Section 5. Finally, I assume that the distribution of the flow value of leisure/home production, F, is Log-Normal with parameters μ_κ and σ_κ, both of which I calibrate to match moments in the data below.

4.1.2 Parameter values: DMP model

I begin by considering the model with no monitoring (i.e., $\mathring{m} = 0$). I calibrate all of the parameters of this model to match values from existing literature or moments from before 2000.

[26] The parameter $\bar{\xi} \equiv \max\{\xi_{jt}\}$ ensures that this function can be interpreted as a proper distribution function. However, from (37) it can be seen that $v_t^n = \eta G(V_t^n) = \eta(V_t^n/\bar{\xi})^{\nu} = \eta\bar{\xi}^{-\nu}(V_t^n)^{\nu}$, implying that this parameter is not separately identified from η. Thus, in Table 2, I report the calibrated value of $\tilde{\eta} \equiv \eta\bar{\xi}^{-\nu}$ rather than η.

Direct calibration. I directly calibrate a subset of the model's parameters to match values commonly used in the literature. The period length is assumed to be one month. Labor productivity is normalized to $y = 1$. I choose a discount factor of $\beta = 0.997$ consistent with a 4% annual steady-state interest rate. I choose a UI replacement rate of $b = 0.25$ following Hall and Milgrom (2008).[27] I choose a vacancy posting cost of $c = 0.17$ following Fujita and Ramey (2012).[28] I choose a value for workers' bargaining power of $\chi = 0.7$ following, e.g., Shimer (2005).

Indirect calibration. I calibrate the remaining standard parameters to match a set of moments from the data and other studies. I choose the match efficiency parameter, ψ, to match the average monthly job-finding probability between 1985 and 2000 of 0.41[29]. I choose the monthly separation rate, δ, to match the average unemployment rate between 1985 and 2000 of 5.9%. I choose the average flow value of leisure/home production, $E[\kappa]$, to yield an average flow value of non-work of 0.71, following Hall and Milgrom (2008) (which is also approximately the midpoint of the range of values of the opportunity cost of employment identified by Chodorow-Reich and Karabarbounis (2016)).[30] I choose the parameter scaling the entry process, $\tilde{\eta}$, to match the average vacancy duration of 21 days identified by Barron et al. (1997) (and later used by Fujita and Ramey (2012)).[31]

Job creation elasticity. The final parameter in the model without monitoring is the elasticity parameter of the job creation process, ν. Because this parameter is not entirely standard in DMP-style models (that have traditionally focused on the limiting case of free entry) and because of its importance for the existence of multiplicity (a point that I return to below in Section 4.3.2), it bears briefly discussing how this parameter is calibrated. A natural approach is to target the wage elasticity of labor demand, an object that has been estimated by a large and diverse literature. Hamermesh (1993) provides an early survey of this literature, finding values in a range around -0.3. Dube (2019) surveys the minimum wage literature, which is also interested in this elasticity, and finds a median estimate of the own-wage elasticity of employment of -0.17. Yet another approach is that of Beaudry et al. (2013, 2018), who estimate city level and city-industry level wage elasticities of labor demand of -0.3 and -1, respectively. In view of these studies, as a baseline, I calibrate ν to generate a model-implied wage elasticity of labor demand of -0.3. In Section 4.3.2 below, I explore the sensitivity of the results to this choice.

4.1.3 Parameter values: Monitoring

In the second step, I turn to the version of the model with monitoring that I will use throughout the remainder of the paper, while keeping parameters discussed above fixed at their calibrated

[27] Hall and Milgrom (2008) argue that this value lies in the middle of the range of plausible values identified in the literature.

[28] Fujita and Ramey (2012) arrive at this value based on a calculation using data from Barron et al. (1997).

[29] I compute the average monthly job-finding probability using data on the total number of unemployed workers (u_t) and the number of short-term unemployed workers (u_t^{ST}) via $p_t = 1 - (u_{t+1} - u_{t+1}^{ST})/u_t$, following Shimer (2012).

[30] In the model with monitoring, the *full distribution* of leisure/home production values, F, is important. In the model without monitoring, on the other hand, only the *first moment* of that distribution, $E[\kappa]$, is important, since, absent an ability to monitor, all workers are assumed to enjoy leisure/home production and search only at the end of the period. Thus, because F is Log-Normal (and so has two parameters, μ_κ and σ_κ), the single restriction that the first moment of this distribution (plus UI transfers) match the average flow value of non-employment from Hall and Milgrom (2008) is consistent with a continuum of possible combinations of the two parameters μ_κ and σ_κ. The second step of the calibration procedure, in which I consider the model with monitoring, serves to uniquely identify both of these parameters.

[31] See Footnote 26 for a discussion of why I calibrate $\tilde{\eta}$ rather than η.

values. Since the full *distribution* of leisure/home production values, F, is critical for determining the equilibrium monitoring cutoffs, I use this version of the model, together with two restrictions, to pin down the parameters of the distribution, μ_κ and σ_κ. Specifically, I choose μ_κ and σ_κ so that the model with monitoring (i) continues to yield an average flow value of non-work of 0.71 when no workers monitor, and (ii) implies that 7% of workers monitor the arrival of vacancies in the low-monitoring steady state.[32] The first restriction simply requires that the model with monitoring that I will analyze be consistent with the calibration of the model without monitoring above (see Footnote 30 for discussion of this point). The second restriction is based on values identified by Stevenson (2009), who studies online job search using the Current Population Survey's (CPS) Computer and Internet Use Supplement from 1998, 2001 and 2003.[33]

Table 2 summarizes parameter values. See Appendix C for complete details and discussion.

Table 2: Calibration

Concept	Parameter	Value
Discount factor	β	0.997
Separation rate	δ	0.03
UI replacement rate	b	0.25
Leisure/home production (mean)	μ_κ	-0.77
Leisure/home production (std. dev.)	σ_κ	0.20
Productivity	y	1
Bargaining parameter (labor's share)	χ	0.7
Vacancy posting cost	c	0.17
Job creation (scale)	$\tilde{\eta} \equiv \eta\bar{\xi}^{-\nu}$	0.02
Job creation (elasticity)	ν	0.01
Matching efficiency	ψ	0.81

Notes: Monthly frequency. Moments computed by the author for calibration based on data between 1985 and 2000.

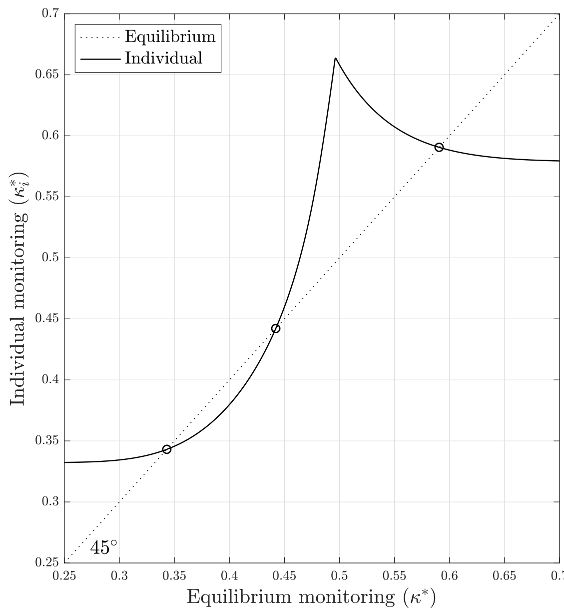
4.2 Steady-state equilibria

The steady-state equilibria of the full dynamic model cannot be solved for analytically, so I proceed numerically. Figure 2 plots the (steady-state) best-response function for worker i in equation (34). As usual, the intersection of the best-response function with the 45-degree line identifies the

[32] I thus assume that the monitoring economy began in the low-monitoring steady state in the early years in which monitoring technologies were available. One way to interpret this assumption is that technological novelty (such as would have been associated with job search platforms in the early 2000s) serves to coordinate individuals' beliefs on the low-engagement (in this case, low-monitoring) equilibrium.

[33] Specifically, Stevenson (2009) finds that (i) averaging across 1998, 2001 and 2003, roughly 27% of unemployed workers and 3% of non-participants searched for jobs online, (ii) among non-employed job seekers using the internet for search, roughly half were unemployed and the other half were non-participants, and (iii) roughly half of online searchers reported that job boards were their primary way of finding work online (as opposed to using general search engines or company websites). Taken together, these values imply that approximately 7% of non-employed workers used online job boards to find work in the early 2000s. I view this as the best available empirical proxy for the extent of monitoring in the early years in which such technologies were available. Note that I focus on both unemployed workers and non-participants, as it is widely recognized that both classes of workers search and transition into employment despite the different official designation.

Figure 2: Best-response function



Notes: The figure plots the best-response function for a typical worker (solid line) and the condition for symmetric equilibrium (dashed 45-degree line). Steady states correspond to the intersections of the two lines.

equilibria of the monitoring game (i.e., where $\kappa_{it}^* = \kappa_t^*$) and thus the steady states of the model. As may be seen in Figure 2, there are three steady states. Table 3 reports the implied values of the targeted moments for all three steady states. It also reports the implied values of several other moments of interest including the wage (w), job creation (v^n/v),[34] the job-finding probability for monitoring workers (p^m), and the job-finding probability for non-monitoring workers (p^w).

The steady states with higher levels of monitoring feature (i) higher average job-finding rates, (ii) lower unemployment, (iii) lower wages, and (iv) more job creation. These observations—in particular, the positive correlation between wages and unemployment across steady states—reflect endogenously different levels of monitoring across steady states, sustained by workers' self-confirming beliefs about the actions of other workers in the economy. Specifically, when workers believe that many other workers are monitoring, they know that there will be relatively few jobs available in the aftermarket. This, in turn, makes forgoing leisure/home production in order to monitor necessary to avoid falling to the back of the queue for jobs and remaining unemployed. Because monitoring is frictionless whereas the aftermarket is not, this leads to higher job-finding rates and so lower unemployment. Moreover, because monitoring requires forgoing leisure/home production, when a large fraction of workers monitor, unemployment looks less attractive in expectation (because it is likely a worker will find monitoring necessary in the event of job loss), which in turn depresses wages, thus stimulating entry and job creation. If job creation is not too elastic, an issue I discuss in more detail below, then this response in job creation is not so strong as to offset the job-depleting effects of high monitoring, and the high-monitoring equilibrium is sustained. This logic can be reversed to understand the coexistence of a low-monitoring equilibrium.

[34] I measure job creation as the share of all vacancies that are new entrants: v^n/v.

Table 3: Description of equilibria

	Data	DMP model	Monitoring model		
		$\kappa^* = 0$	*Low-*κ^*	*Int.-*κ^*	*High-*κ^*
Targeted moments					
Unemployment rate (u)	$\boxed{5.9\%}$	$\boxed{5.9\%}$	5.5%	4.3%	3.6%
Job-finding prob. (p)	$\boxed{0.41}$	$\boxed{0.41}$	0.44	0.58	0.69
Average vac. duration	$\boxed{\text{21 days}}$	$\boxed{\text{21 days}}$	17 days	8 days	0 days
Share monitoring ($F(\kappa^*)$)	$\boxed{7\%}$	0%	$\boxed{7\%}$	42%	89%
Untargeted moments					
Wage (w)		0.96	0.95	0.91	0.83
Job creation (v^n/v)		0.80	0.84	0.98	1
Job-finding prob. (p^m)		n/a	1	1	0.78
Job-finding prob. (p^w)		n/a	0.40	0.28	0

Notes: Boxed moments were targeted in calibration. See Section 4 and Appendix C for further details.

4.3 Alternative models

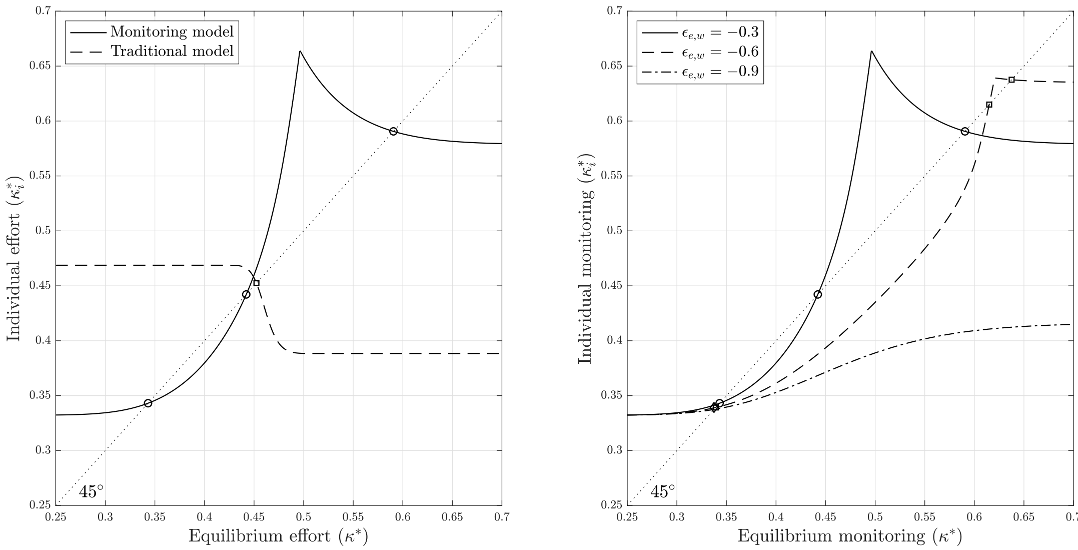
The two critical features of the model that I study are (i) conceptualizing search as a monitoring decision and (ii) a relatively inelastic entry process. Figure 3 depicts the steady-state best-response functions in alternative models in which these features are not present. Specifically, the dashed line in Figure 3a depicts a model with a traditional interpretation of search but the same entry elasticity, while the dashed lines in Figure 3b depict the monitoring model with more elastic entry processes.

4.3.1 Traditional search

As discussed in Section 2, the traditional approach to modeling search decisions in quantitative macroeconomic models is to assume that searchers and non-searchers are perfect substitutes in the match function. In this interpretation, non-searchers are simply ineffective searchers whose match rate is proportional to that of searchers (as in, e.g., (15) and (16) in Section 2). Recent examples of such models include Krusell et al. (2017) and Cairo et al. (2021). In Appendix B, I formally consider such a model. The dashed line in Figure 3a depicts the best-response function implied by embedding this more traditional model of search within an otherwise identical framework and calibration procedure to the one I describe above.

Consistent with the intuition from Section 2, the best-response function is decreasing, implying that this more traditional model has a unique steady state. This observation is important because it highlights that a key feature of the model giving rise to multiple equilibria is the conceptualization of search as a monitoring decision.

Figure 3: Alternative models



(a) Traditional search (b) Elastic entry

Notes: Panel (a) depicts the best-response function in the baseline model (solid line) and the best-response function in a model with a traditional interpretation of search (dashed line). Panel (b) depicts the best-response function in the baseline model (solid line) and the best-response functions for two otherwise identical models with more elastic job-creation processes (dashed lines). In all cases, steady states are depicted by the intersections of the best-response functions with the 45-degree line.

4.3.2 Elastic entry

As discussed both in the Introduction and in Section 4.1 above, the existence of multiplicity depends critically on the elasticity of the job-creation process. The reason for this is the following: As more workers choose to monitor the arrival of vacancies, the expected value of unemployment falls as workers come to expect that, if they become (or remain) unemployed, they will likely need to forgo leisure/home production in order to monitor vacancies to avoid falling to the back of the queue. A lower expected value of unemployment worsens workers' outside option in the wage bargain and thus depresses wages. When job creation is highly elastic with respect to the wage, falling wages associated with high levels of monitoring cause entrepreneurs to create a large number of new jobs, which alleviates congestion among those workers at the back of the queue choosing not to monitor, thus reducing the urgency of monitoring and mitigating the increasing returns that enable multiplicity when job creation is relatively unresponsive to wages.

The dashed line in Figure 3b depicts the best-response functions in the monitoring model when that model is calibrated to generate a wage elasticity of labor demand of $\epsilon_{e,w} = -0.6$ rather than the baseline of $\epsilon_{e,w} = -0.3$. The dashed-dotted line depicts the best-response function with $\epsilon_{e,w} = -0.9$. The figure illustrates that increasing the elasticity of job creation tends to flatten out the best-response function, consistent with the intuition in the preceding paragraph. Notwithstanding this, multiplicity continues to arise when this elasticity is doubled (in absolute value), but when it becomes sufficiently large, as in the dashed-dotted line, multiplicity is eliminated.

4.4 Welfare

The differences across equilibria reflected in Table 3 translate into differences in social welfare. To understand the forces that influence social welfare in the model, it is useful to consider the steady state of the model as a function of the monitoring cutoff κ^*, that is, treating κ^* as exogenous and not imposing (34) (as in Figures 2 and 3). Social welfare, expressed as a function of κ^*, is then:

$$\Omega(\kappa^*) = \underbrace{(1-u(\kappa^*))y}_{\text{Production } (\Omega_y)} + \underbrace{u(\kappa^*)\int_{\kappa^*}^{\infty}\kappa f(\kappa)d\kappa}_{\text{Leisure/home prod. } (\Omega_l)} - \underbrace{\left[cv(\kappa^*) + \tilde{\eta}\frac{\nu}{1+\nu}\left(V^n(\kappa^*)\right)^{\nu+1}\right]}_{\text{Vacancy costs + Entry costs } (\Omega_v)}. \tag{47}$$

The first term (Ω_y) corresponds to total output from employed workers matched with entrepreneurs. The second term (Ω_l) corresponds to the value of leisure/home production among workers who choose not to monitor. The third term (Ω_v) corresponds to the sum of the costs of maintaining a posting for unmatched vacancies and the entry costs for entrepreneurs.[35]

The first term is increasing in κ^* because, as may be seen in Table 3, unemployment is decreasing in κ^*. The second term, on the other hand, is decreasing in κ^*: More workers choosing to monitor implies both (i) fewer workers unemployed who can enjoy leisure/home production and (ii) a smaller *share* of unemployed workers choosing to enjoy leisure/home production. The final term is less clear: On the one hand, more workers monitoring means that there are fewer vacancies that are not matched immediately, and thus fewer firms paying costs of maintaining a vacancy (c). On the other hand, more monitoring depresses wages by reducing the (ex ante) value of unemployment, which in turn increases the value to an entrepreneur of posting a new vacancy. This results in more entry from entrepreneurs with higher costs, which detracts from social welfare.

Quantitatively, the important tradeoff for social welfare is between the first two terms, Ω_y and Ω_l. Moreover, the first term can be decomposed into two components, reflecting the channels through which more monitoring increases total production: A component due to reduced matching frictions associated with monitoring (recall that the monitoring phase is frictionless whereas the aftermarket is not, so as more workers decide to monitor, the matching process becomes more efficient), and a component due to more entry/job creation resulting from monitoring. Formally, making explicit the dependence of Ω_y on κ^* via unemployment, which in turn depends on κ^* both directly and via entry of new firms, Ω_y can be written as

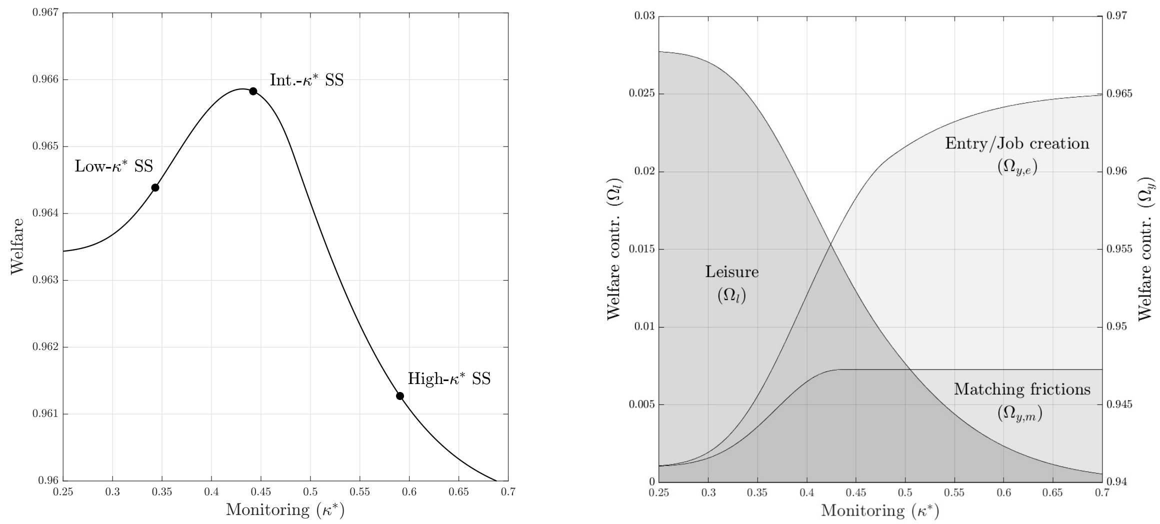
$$\Omega_y(\kappa^*) = \Big(1 - u(\kappa^*, v^n(\kappa^*))\Big)y \tag{48}$$

$$= \underbrace{\Big(1 - u(\kappa^*, \underline{v}^n)\Big)y}_{\text{Matching frictions } (\Omega_{y,m})} + \underbrace{\Big(u(\kappa^*, \underline{v}^n) - u(\kappa^*, v^n(\kappa^*))\Big)y}_{\text{Entry/job creation } (\Omega_{y,e})} \tag{49}$$

where $\underline{v}^n$ is some fixed level of job creation. Fixing $\underline{v}^n = 0.25$, Figure 4a plots total social welfare Ω as a function of κ^*, indicating the levels of monitoring associated with the three steady states. Figure 4b plots the three most important components of social welfare, $\Omega_{y,m}$, $\Omega_{y,e}$, and Ω_l, also as functions of κ^*, with the left axis measuring the contribution of leisure/home production and the right axis measuring the contribution of the components of output. Figure 4 reveals several

[35] By assumption the government runs a balanced budget (see equation (44)), implying that b is a pure transfer

Figure 4: Social welfare



(a) Total

(b) Components

Notes: Panel (a) depicts social welfare (Ω) as a function of the equilibrium monitoring cutoff (κ^*). Panel (b) depicts various components of social welfare (Ω_l, $\Omega_{y,e}$, and $\Omega_{y,m}$) as functions of the equilibrium monitoring cutoff (κ^*).

interesting features of how monitoring influences social welfare. First, as can be seen in Figure 4a, social welfare is non-monotone with respect to the monitoring cutoff κ^*. This reflects the competing forces of leisure and output described above. Second, the low-monitoring steady state Pareto dominates the high-monitoring steady state. As can be seen in Figure 4b, this results from the fact that, at sufficiently high levels of monitoring, the increased matching efficiency and increased job creation associated with more monitoring contribute an increasingly small amount to social welfare, whereas the leisure/home production costs of more monitoring are continuing to substantially detract from social welfare. Interestingly, this implies that the model reverses the Pareto ranking obtained in Diamond (1982a), in which the "high activity" equilibrium necessarily dominates the "low activity" one. This important difference reflects the strong congestion externalities associated with monitoring technology that are not internalized by job seekers who are trying to avoid being shut out of the queue for jobs. This feature of the model, which results from the inelastic job creation process, is similar to the "rat race" among workers described in Landais et al. (2018) and Toohey (2021). Finally, the model implies that the gains from increased monitoring are largely due to entry rather than reduced matching frictions.

It should be emphasized that these results are relatively sensitive to the calibration of the model. Nevertheless, that the model can, in principal, overturn the basic normative implication of Diamond (1982a), and indeed does (at least when comparing the high- and low-monitoring steady states) is an important implication of the analysis to which I will return in the discussion of the Great Recession in Section 5.

and thus does not appear in (47).

5 Dynamics and Quantitative Implications

I next study the model's global dynamics and explore some quantitative implications. Specifically, I focus on three central implications for aggregate fluctuations in the labor market. First, I show numerically that, for a wide range of initial conditions, the economy has equilibria converging to different steady states. This implies that the economy is susceptible to belief-driven fluctuations in monitoring and thus labor supply. Second, I show that under a natural refinement of the set of equilibrium trajectories, a demand shock that causes a sufficiently large increase in unemployment forces the economy onto a unique path towards the high-monitoring steady state, whereas multiplicity remains for small shocks, implying that the proposed refinement both enables large transitory exogenous shocks to permanently influence the economy while still allowing a role for purely endogenous belief-driven fluctuations. Third, I argue that the model can help to explain several features of the recovery from the Great Recession that are difficult to explain with traditional models.

5.1 Frictionless economy ($\psi \to \infty$)

To facilitate exposition of the model's dynamics, I focus on the limiting case in which $\psi \to \infty$. This corresponds to a frictionless aftermarket, implying that $\hat{q}_t = q_t^n = q_t^o = q_t = 1$ and thus that $v_t = v_t^n$. It follows that, in this case, the model has a single endogenous state variable (u_t) and two forward-looking variables (v_t and κ_t^*), which significantly facilitates analysis and graphical exposition of the model's global dynamics. Specifically, in Appendix D, I show that the frictionless model's dynamics are characterized by three equations:

$$u_{t+1} = u_t + \delta(1-u_t) - (1-\delta)v_t \tag{50}$$

$$v_t = \eta G\left(-c + \beta(1-\delta)\left[y_{t+1} - \omega(\kappa_{t+1}^*, u_{t+1}, v_{t+1})\left(\tfrac{1-(1-b)u_{t+1}}{1-u_{t+1}}\right) + c + G^{-1}\left(\frac{v_{t+1}}{\eta}\right)\right]\right) \tag{51}$$

$$\kappa_t^* = \begin{cases} \beta(1-\delta)\left(\frac{1-v_t/u_t}{1-F(\kappa_t^*)}\right)\frac{\chi}{1-\chi}\left(y_{t+1} - \omega(\kappa_{t+1}^*, u_{t+1}, v_{t+1})\left(\frac{1-(1-b)u_{t+1}}{1-u_{t+1}}\right) + c\right) & \text{if} \quad F(\kappa^*) \leq v_t/u_t \\ \beta(1-\delta)\left(\frac{v_t/u_t}{F(\kappa_t^*)}\right)\frac{\chi}{1-\chi}\left(y_{t+1} - \omega(\kappa_{t+1}^*, u_{t+1}, v_{t+1})\left(\frac{1-(1-b)u_{t+1}}{1-u_{t+1}}\right) + c\right) & \text{if} \quad F(\kappa^*) > v_t/u_t \end{cases} \tag{52}$$

where

$$\omega(\kappa_t^*, u_t, v_t) = \begin{cases} \dfrac{\frac{\chi}{1-\chi}(y_t+c) + \int_{\kappa_t^*}^{\infty}\kappa dF(\kappa) - \kappa_t^*(1-F(\kappa_t^*))}{1-b+\frac{\chi}{1-\chi}\left(\frac{1-(1-b)u_t}{1-u_t}\right)} & \text{if} \quad F(\kappa^*) \leq v_t/u_t \\ \dfrac{\frac{\chi}{1-\chi}(y_t+c) + \int_{\kappa_t^*}^{\infty}\kappa dF(\kappa) - \left(\frac{u_t-v_t}{v_t}\right)\kappa_t^* F(\kappa_t^*)}{1-b+\frac{\chi}{1-\chi}\left(\frac{1-(1-b)u_t}{1-u_t}\right)} & \text{if} \quad F(\kappa^*) > v_t/u_t. \end{cases} \tag{53}$$

The frictionless system characterized by equations (50)-(53) continues to have three steady-state equilibria, as in the full quantitative model described in Section 3. Furthermore, these three steady states retain the same qualitative features of those in Section 3. I thus proceed with an analysis of the frictionless model.

5.2 Dynamics

With equations (50)-(53) in hand, I am prepared to study the model's dynamic properties. I first analyze the model's local dynamics in a neighborhood around each of the three steady-state equilibria. I then proceed to analyze the model's global dynamics away from these steady states.

5.2.1 Local dynamics

To analyze the model's local dynamics, I linearize the system in (50)-(53) around each of the three steady states. The linearized system of equations is given by

$$\begin{bmatrix} u_{t+1} \\ v_{t+1} \\ \kappa^*_{t+1} \end{bmatrix} = \begin{bmatrix} u \\ v \\ \kappa^* \end{bmatrix} + \begin{bmatrix} \frac{\partial u_{t+1}}{\partial u_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial u_{t+1}}{\partial v_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial u_{t+1}}{\partial \kappa^*_t}\big|_{\{u,v,\kappa^*\}} \\ \frac{\partial v_{t+1}}{\partial u_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial v_{t+1}}{\partial v_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial v_{t+1}}{\partial \kappa^*_t}\big|_{\{u,v,\kappa^*\}} \\ \frac{\partial \kappa^*_{t+1}}{\partial u_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial \kappa^*_{t+1}}{\partial v_t}\big|_{\{u,v,\kappa^*\}} & \frac{\partial \kappa^*_{t+1}}{\partial \kappa^*_t}\big|_{\{u,v,\kappa^*\}} \end{bmatrix} \begin{bmatrix} u_t - u \\ v_t - v \\ \kappa^*_t - \kappa^* \end{bmatrix} \tag{54}$$

where variables without time subscripts indicate steady-state values. The Jacobian in (54) determines the system's local dynamics around each of the three steady states. Because it is not possible to solve for the elements of the Jacobian analytically, I do so numerically. Based on the parameterization in Section 4 (with $\psi \to \infty$), the Jacobian in (54) has two eigenvalues outside of the unit circle and one eigenvalue inside of the unit circle for all three steady states. Because the model has one predetermined variable (u_t) and two forward-looking variables (v_t and κ^*_t), this implies that a unique non-explosive solution exists for all three steady states. Momentarily neglecting multiplicity of steady states, this result implies that, given some initial condition u_0 in a neighborhood of any of the steady states, κ^*_t and v_t will immediately jump to be on the saddle path that leads the economy back to the steady state.

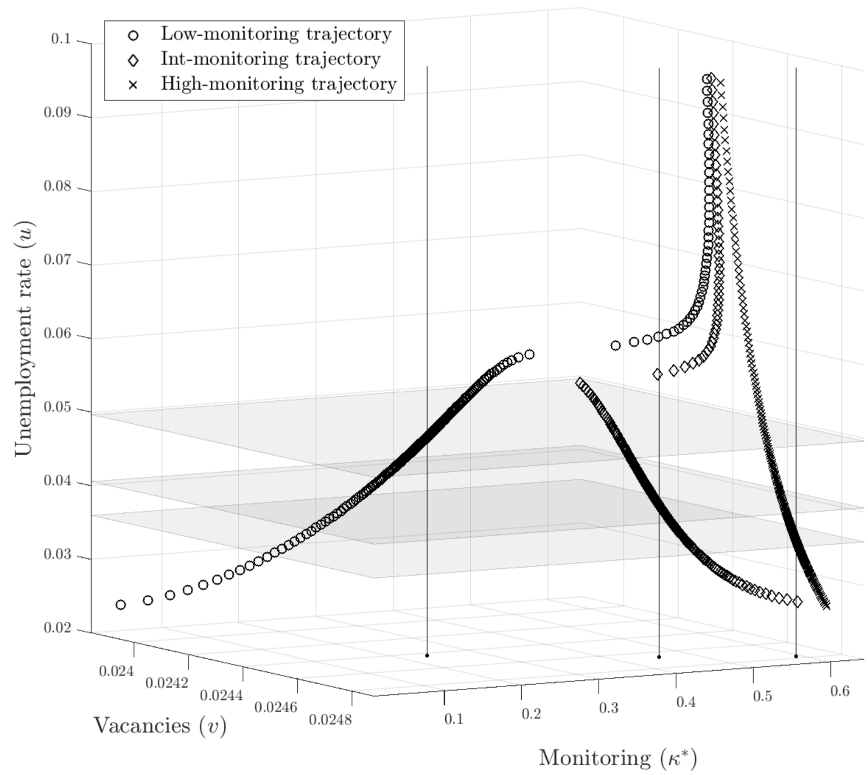
5.2.2 Global dynamics

Knowledge of the model's local dynamics as discussed above can be used to construct global stable manifolds and thus to study the model's global dynamics. Specifically, following Judd (1999) and Brunner and Strulik (2002), I numerically construct the global stable manifolds converging to the three steady states using backward integration. In effect, this method takes advantage of the fact that, when the model is solved in reverse time, the stable manifolds in forward time become the unstable manifolds. Paths attracted to the unstable manifold thus feature "exponential tracking"—i.e., exponentially decreasing deviations from the stable manifold as the system is run in reverse time. This implies that the global stable manifolds will be well-approximated by beginning from a point in a neighborhood of each steady state that need not lie precisely on the (zero measure) global stable manifold and running the system in reverse time.[36]

Figure 5 plots the three global stable manifolds computed in this way. The grey planes depict the three steady-state levels of unemployment. The vertical lines depict the three steady-state pairs of values $\{\kappa^*, v\}$. Figure 5 highlights a fact about the model's global dynamics that will be essential for the analysis that follows: For a wide range of initial conditions (which can be visualized as planes at fixed values of u_t, such as those associated with the three steady-state values of u_t), there

[36] See Atolia and Buffie (2009) for a good discussion of this point and solution methods for models with two or more state variables. As they point out, analysis of global dynamics is considerably easier with a single state variable, which is part of the reason why I prefer to consider the frictionless model for the quantitative analysis in this section.

Figure 5: Global dynamics



Notes: The figure depicts the global stable manifolds converging towards the three steady states. The shaded planes correspond to the steady-state levels of unemployment. The vertical lines depict the three steady-state pairs of values $\{\kappa^*, v\}$. Global dynamics are non-unique at any value of u_t through which multiple stable manifolds pass.

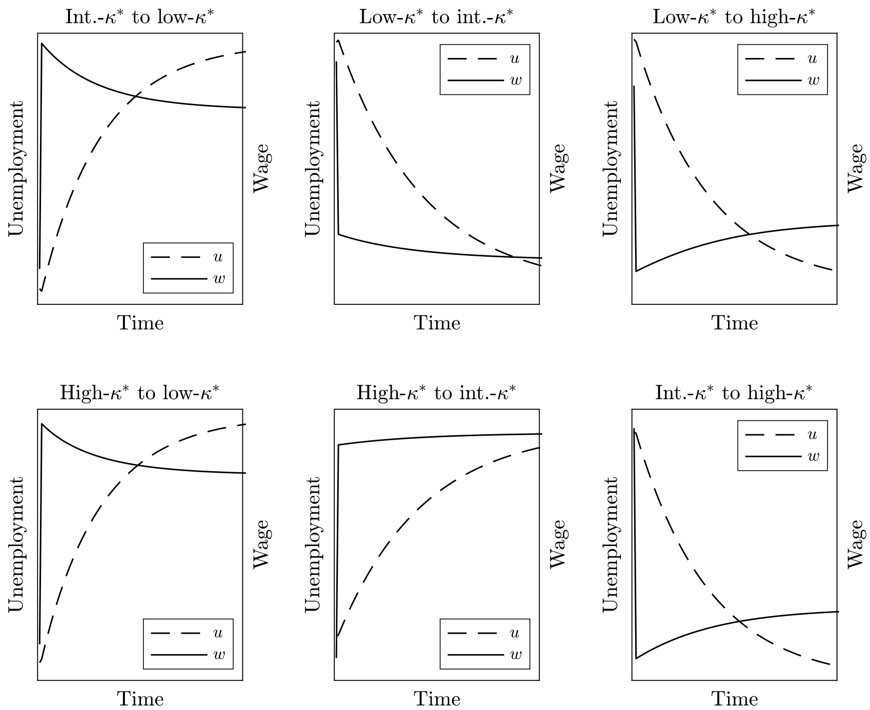
are multiple equilibria leading to different steady states. I discuss some implications of this result below.

5.3 Belief-driven labor supply fluctuations

In Section 4, I documented that the economy has multiple steady-state equilibria (which continues to be the case as $\psi \to \infty$). Figure 5 illustrates that the economy also has multiple *dynamic* equilibria. More precisely, there is a range of initial conditions for the economy, i.e., $u_0 \in [\underline{u}, \bar{u}]$, for which the economy can converge to any of the three steady states. Indeed, inspection of Figure 5 reveals that this is possible from any of the steady states and for most of the depicted range of levels of unemployment. Within this range, then, which of the three paths is actually realized is not determined by fundamentals, but rather by workers' beliefs about the intensity with which other workers in the economy are monitoring the arrival of vacancies.

Functionally, the multiplicity of dynamic equilibria observed in Figure 5 implies that the economy is susceptible to autonomous shocks to workers' beliefs that can permanently alter the trajectory

Figure 6: Belief-driven labor supply shocks



Notes: The panels depict impulse responses of unemployment (dashed lines) and wages (solid lines) associated with the six possible belief shocks. In each case, the correlation between wages and unemployment over short horizons following the shocks are positive, consistent with the interpretation of the belief shocks as labor supply shocks.

of the economy by shifting the economy from a path converging to one steady state to a path converging to another. Because such changes in beliefs manifest through unemployed workers' decisions to actively monitor the arrival of vacancies, the model can be understood as providing a theory of endogenous belief-induced labor supply fluctuations.[37] Specifically, Figure 6 depicts the responses of wages (solid lines) and unemployment (dashed lines) to the six possible belief shocks shifting the economy between the trajectories depicted in Figure 5.[38] In each case, I assume that the economy begins at a particular steady state and then workers' beliefs suddenly (and unexpectedly) change such that the economy is expected to converge to a different steady state. The first column corresponds to the two shocks that shift the economy to the low-monitoring steady state, the second column corresponds to the two shocks that shift the economy to the intermediate-monitoring steady state, and the third column corresponds to the two shocks that shift the economy to the high-monitoring steady state.

This figure is important for two reasons. First, it illustrates the rich dynamics that can be induced

[37] As I have described it above, the model can accommodate *unanticipated* shocks to beliefs without any additional analysis. Such shocks can be thought of as simply shifting the economy between the various global stable manifolds. This is the approach taken by Eeckhout and Lindenlaub (2018). Incorporating *anticipated* shocks to beliefs would require modeling beliefs as following, e.g., a Markov process, as in Kaplan and Menzio (2016).

[38] There are six possible shocks because there are three steady states and from each steady state belief shocks can shift the economy to a path converging towards either of the two other steady states.

by shocks to workers' beliefs. Second, it illustrates why these shocks can be interpreted as labor supply shocks: In all six cases, the effect of the shock in the first several periods induces a *positive correlation* between wages and unemployment. For example, in the top-left panel (corresponding to a shock that shifts the economy from the intermediate-monitoring steady state to the low-monitoring steady state), the shock raises wages substantially on impact and, while unemployment is predetermined, after the period of the shock, unemployment begins to rise, inducing a positive correlation. Eventually, in the absence of further shocks, the fact that wages fall after their initial rise will come to dominate and render the correlation negative, but this takes a number of periods to play out because the initial shock to wages is so large. This observation that belief shocks induce a positive short-run correlation between wages and unemployment, which holds for all six possible shocks, will be critical for motivating my approach to identifying labor supply shocks, based on short-run restrictions on the relationship between unemployment and wages, in Section 6.[39]

5.4 The Great Recession

I next study the model's implications for the Great Recession and the subsequent recovery. After suggesting a refinement of the set of equilibria to help pin down the model's dynamics, I show that a transitory shock to the entrepreneurial sector in the model can force the economy onto a trajectory leading to the high-monitoring steady state. This observation has two important implications: First, it implies that transitory demand shocks can permanently shift labor supply. Second, and for this reason, it can help to account for several features of the recovery from the Great Recession that are difficult to understand in the context of purely demand-side explanations.

5.4.1 Equilibrium selection: Non-reversing return to monitoring

While the existence of multiple equilibria leads to interesting and potentially important theoretical insights, such as the possibility of belief-driven fluctuations in labor supply as described above, multiplicity also presents an empirical challenge since the model does not make unique predictions about the trajectory of the economy following, e.g., a disturbance to fundamentals. Obtaining such predictions, that can then be compared with the data, requires a systematic way of refining the set of candidate equilibria. Below, I propose a simple refinement based on a restriction on workers' beliefs along the equilibrium path, and use it to study the model's response to a transitory demand shock calibrated to emulate the Great Recession.

Definition 2 (Non-reversing return to monitoring criterion). *An equilibrium satisfies the non-reversing return to monitoring criterion if workers believe that, in the absence of fundamental shocks, along the equilibrium path:*

(i) The economy will stay on a single global stable manifold;

(ii) The return to monitoring ($p_t^m - p_t^w$) will never switch between being increasing and decreasing. That is, either $\mathbb{E}_t\left[\frac{\partial(p_{t+\tau}^m - p_{t+\tau}^w)}{\partial \kappa_{t+\tau}^*}\right] > 0$ for all $\tau > 0$ or $\mathbb{E}_t\left[\frac{\partial(p_{t+\tau}^m - p_{t+\tau}^w)}{\partial \kappa_{t+\tau}^*}\right] < 0$ for all $\tau > 0$.

The first restriction simply rules out workers anticipating shifts between the trajectories depicted in Figure 5. In effect, this restricts the set of equilibrium paths to the trajectories depicted in the figure. The second restriction, in turn, provides a selection criterion among those paths: It rules

[39] Readers wishing to be convinced of this point can skip ahead to Figure 9 in Section 6, where I plot the correlation between unemployment and wages at various horizons following the six belief shocks.

out workers anticipating the equilibrium transiting between one in which the return to monitoring effort is positive and one in which it is negative (or vice versa). Importantly, the criterion above does not rule out the possibility of fundamental shocks shifting the economy between the global stable manifolds; rather it states that the set of global stable manifolds that the economy can reach from any initial position is restricted based on workers' beliefs about the evolution of the economy.

The criterion proposed in Definition 2, and specifically the second part of the criterion, is appealing for two reasons. First, it can be understood as a restriction on the "sophistication" of workers' beliefs: Trajectories in which the qualitative nature of a choice fundamentally changes such as would be implied by a reversal of the return to monitoring are inherently more complicated than trajectories in which such a change does not occur. Second, it is relatively weak: In particular, it does not rule out multiplicity in large regions of the state space, which implies that it does not rule out the possibility of endogenous shocks to beliefs such as those described in Section 5.3. Importantly, however, it can function to *eliminate* multiplicity when a shock raises unemployment sufficiently, a result that is reminiscent of Proposition 1. The basic reason for this is that a sufficiently high unemployment rate forces the economy into a position in which newly posted vacancies are the short end of the market in the monitoring phase—a situation in which the return to monitoring is *negative* (see the bottom line of equation (9)). When this is the case, in order for the economy to converge to either the low- or intermediate-monitoring steady states, it must transit into a position in which workers eventually become the short end of the market in the monitoring phase—a transition which necessitates a reversal such that the return to monitoring eventually becomes *positive* (see the top line of equation (9)). This transition can be seen in the discrete jumps observed in the low- and intermediate-monitoring trajectories as the economy returns from a high unemployment rate in Figure 5. Thus, when a shock is sufficiently large to drive the economy to a high level of unemployment, the only remaining equilibrium is the one that leads to the high-monitoring steady state. In what follows, I explore the implications of this observation for understanding the Great Recession.

5.4.2 Financial shocks and the Great Recession

To study the model's implications for the Great Recession and the economy's subsequent recovery, I study a shock to the cost of financing investments in job creation among entrepreneurs, $\tilde{\eta}$.[40] While this is not intended to be a quantitative exercise, it is a simple way to qualitatively explain several facts about the recovery from the Great Recession following an impulse that can be understood as emanating from the financial sector.

Formally, in the context of the model described above, I assume that $\tilde{\eta}$ follows an AR(1) process,

$$\tilde{\eta}_{t+1} = \rho\tilde{\eta}_t + (1-\rho)\tilde{\eta} \tag{55}$$

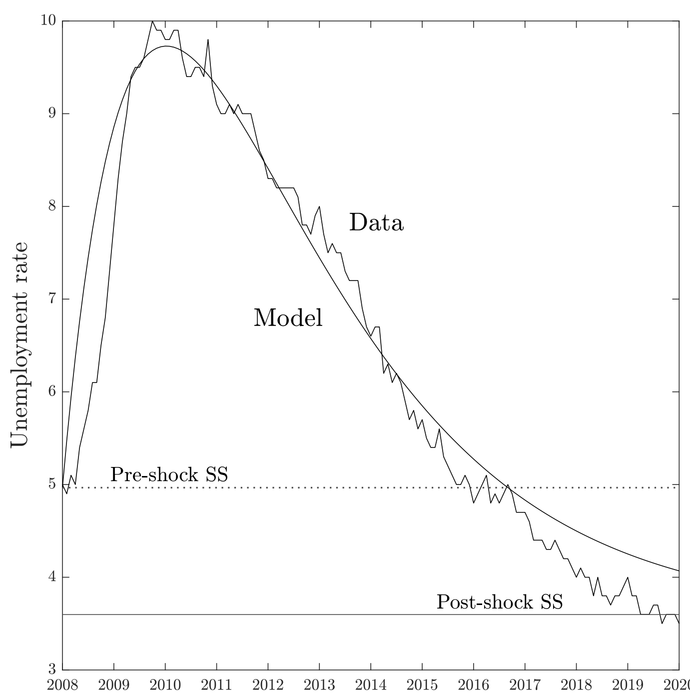
where $\tilde{\eta}$ is the steady-state value calibrated in Section 4. Agents in the model are assumed to know the AR(1) process but do not anticipate shocks to $\tilde{\eta}_t$. I set $\rho = 0.95$ and consider a 25% contraction in $\tilde{\eta}_t$, values chosen to roughly match the peak unemployment rate of 10% in 2009 and the unemployment rate of 5% in 2016 when it had fully returned to its pre-recession average. Im-

[40] Recall from Section 3 that, in order to undertake an investment project and thus create a job, an entrepreneur must pay a sunk investment cost drawn from distribution G with support $[0, \bar{\xi}]$. Alternatively, this can simply be interpreted as a shock to the measure of entrepreneurs in the economy emanating from some other source.

portantly, a shock sufficiently large to generate 10% unemployment eliminates the two equilibrium paths converging to the low- and intermediate-monitoring steady-states under the refinement suggested above. Thus, beginning from the low-monitoring steady-state, the shock forces the economy onto a path converging to the high-monitoring steady state. This observation is important because it implies that, under the equilibrium refinement described above, purely transitory shocks can induce a permanent shift in workers' labor supply decisions. Below, I argue that this fact can help to account for several anomalous features of the recovery from the Great Recession.

Unemployment. Figure 7 depicts the effect of the entrepreneurial shock on the unemployment rate in the model, and compares the model's response with the data. The dashed horizontal line depicts the average unemployment rate prior to the Great Recession, which coincides with the low-monitoring steady state. The solid horizontal line depicts the high-monitoring steady state. The figure is intended to make a simple point: A shock that causes unemployment to rise to 10% causes the economy to jump to an equilibrium trajectory that will ultimately lead to a lower unemployment rate than that from which the economy began. This implication regarding the evolution of unemployment is consonant with the experience of the ten years following the financial crisis and Great Recession, during which unemployment fell to below 4% prior to the Covid-19 crisis. Standard macroeconomic models with unique steady states cannot easily account for this fact.

Figure 7: Unemployment during the Great Recession



Notes: Pre-shock SS corresponds to the (targeted) low-monitoring steady state. Post-shock SS corresponds to the (untargeted) high-monitoring steady state.

Monitoring. What does the model imply about the trajectories of other key labor market variables following the shock? To answer this question, Figure 8 depicts the responses of monitoring effort

(8a), wages (8b), and labor market tightness (8c). Turning first to Figure 8a, we see that monitoring rises substantially on impact of the financial shock, falls slightly for a year or two, and then continues to rise towards its new, higher steady-state level. This monitoring behavior reflects the fact that the shock forces the economy onto a trajectory in which workers believe that other workers are aggressively monitoring new postings and will continue to do so into the future. Because failing to do the same would result in severely depressed job-finding prospects as a result of the temporal segmentation of the matching process in the presence of monitoring technology, the belief that other workers are monitoring aggressively induces individual workers to do the same.[41]

Tightness and wages. On November 5, 2019—almost exactly ten years after unemployment reached 10% in October 2009—John Robertson posted on the Atlanta Fed's MacroBlog about the state of the labor market ten years into the recovery. Robertson writes:

> *Here's a puzzle. Unemployment is at a historically low level, yet nominal wage growth is not even back to prerecession levels (see, for example, the Atlanta Fed's own Wage Growth Tracker). Why is wage growth not higher if the labor market is so tight?*

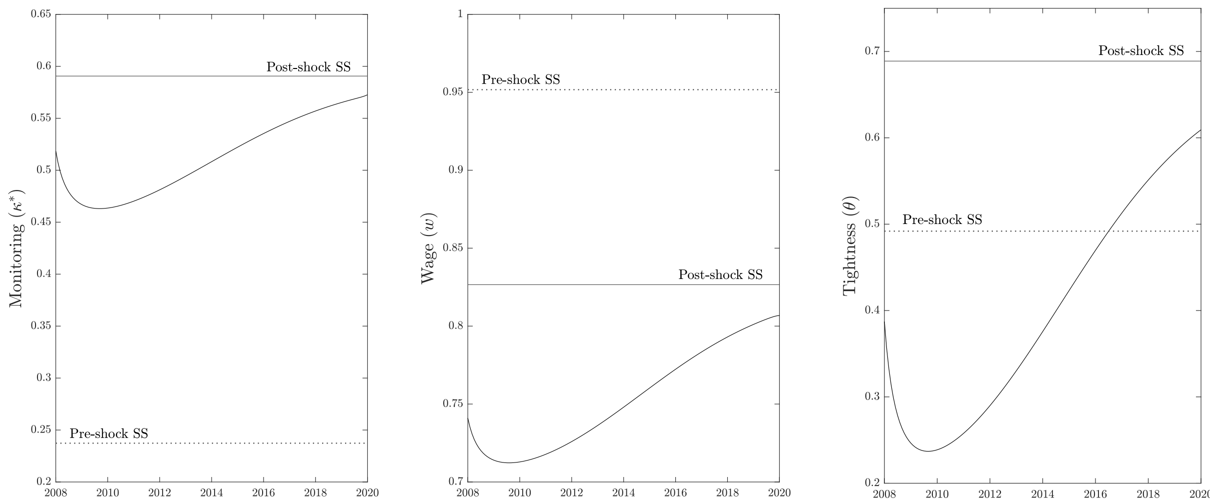
This quote captures two additional features of the recovery from the Great Recession that are perplexing from the standpoint of traditional models: The labor market (ultimately) was quite tight, but wage growth was persistently tepid. In the presence of a large initial shock such as the one described above, the model considered in this paper can qualitatively account for both of these observations. Figures 8b and 8c illustrate this point, plotting the model-implied trajectories of wages and tightness.

Following the shock, the economy is forced onto the high-monitoring trajectory, which *permanently* depresses wages by reducing the expected value of unemployment as workers know they will have to compete aggressively for jobs, and thus forgo leisure/home production, if they return to unemployment. Wages eventually begin to recover, but only to the high-monitoring steady state at which they are lower than they were prior to the shock, as may be seen in Figure 8b. In turn, this weak recovery in wages, coupled with the high level of monitoring, gives rise to a strong recovery in vacancies, which recover to a level that *exceeds* the pre-shock level as entrepreneurs find matches quickly and are able to save on labor costs due to the weak bargaining position of workers. Together with low (and falling) unemployment, this implies that the economy was on a trajectory converging to a historically tight labor market, as may be seen in Figure 8c. Thus, the model can explain the coexistence of tepid wage growth and a historically tight labor market.[42]

[41] While the notion of monitoring considered in this paper is not analogous to "search effort" broadly defined, and thus considerable caution should be taken when comparing the model's implications for monitoring to empirical results concerning search effort, two observations are worth pointing out. First, the large and (more or less) sustained increase in monitoring following the shock would likely lead to measured countercyclicality in monitoring (if such measurement was possible) at business-cycle frequencies based on data from the Great Recession, since unemployment remained elevated for so long. Keeping the above caveat in mind, this is consistent with the result in Mukoyama et al. (2018), who document that search effort is countercyclical based on data largely drawn from the Great Recession period. Second, and more generally, the possibility that large shocks can force the economy between distinct equilibrium paths renders the task of sorting out the cyclicality of variables such as job search a particularly difficult task, which could help to explain why, at least prior to Mukoyama et al. (2018), there seemed to be little consensus on the topic.

[42] Surveys that focus on workers' beliefs about the economy, such as the Survey of Consumer Expectations, do not extend far enough back to make a determination about whether the Great Recession was associated with a permanent shift in workers' beliefs about the job-finding process as implied by the model and discussion above. However, a simple way to measure the beliefs of a *subset* of workers is to use responses to the CPS question about non-participants' reasons for not searching. Responses to this question reveal that the Great Recession coincided with a sharp and

Figure 8: The recovery from the Great Recession



(a) Monitoring (b) Wages (c) Tightness

Notes: Panel (a) depicts the response of equilibrium monitoring intensity (κ^*) to the shock. Panel (b) depicts the response of wages (w) to the shock. Panel (c) depicts the response of labor market tightness ($\theta \equiv v/u$) to the shock. In each case, "Pre-shock SS" refers to the low-monitoring steady state targeted in calibration and "Post-shock SS" refers to the high-monitoring steady state towards which the economy is forced following a sufficiently large impulse.

These observations would be difficult to understand in the context of a traditional model with a unique steady state. Moreover, they would also be difficult to understand in the context of a model of multiple equilibria driven by strategic complementarities in labor demand, since the strong labor demand required to generate a tight labor market would necessarily lead also, and counterfactually, to permanently elevated wages.

6 Some Evidence on Monitoring and its Implications

In this final section, I provide two pieces of evidence in support of the mechanism and its implications. First, using high-frequency panel data on unemployed workers' search behavior, I show that online job search is significantly less lumpy than traditional job search. This is consistent with online job search facilitating monitoring the arrival of new job openings. Second, I show that the contribution of labor supply shocks to aggregate fluctuations increased markedly around the turn of the century. This is consistent with the emergence of endogenous belief-driven fluctuations in labor supply resulting from the rapid growth in online search as discussed above.

seemingly permanent increase in the share of all workers who wanted a job and were available for work who reported that they believed there was no work available and had thus stopped searching. This observation is consistent with the narrative of the Great Recession provided in this paper, wherein the job-finding prospects of workers choosing not to monitor (p^w) would have become permanently depressed following a shock that shifted the economy onto a high-monitoring trajectory, as may be seen in Table 3.

6.1 Micro evidence: The internet makes search less lumpy

The mechanism that I study in this paper is predicated on the idea that online job search technologies reduce the fixed costs of checking for new job listings, thus allowing workers to search frequently—what I refer to as monitoring.[43] This implies that prior to the advent of such technologies, monitoring would have been infeasible, thus rendering individuals' search decisions "lumpy" in the sense of occurring infrequently but for long periods to economize on fixed costs. Evidence that online search is less lumpy than traditional search would thus constitute evidence in support of the central hypothesis of this paper regarding the mode by which technology affects search behavior.

Unfortunately, the vast majority of data on job search is simply too low frequency to make any meaningful determinations about whether online job search is less lumpy than traditional job search at a relevant frequency. Moreover, data on whether individuals are using the internet for job search is both relatively scarce and likely to be contaminated by non-trivial selection effects. To overcome these empirical challenges, I leverage high-frequency panel data from the Survey of Unemployed Workers in New Jersey (SUWNJ) to shed light on the question of whether the internet makes search less lumpy. The SUWNJ is a weekly longitudinal survey (lasting for up to 26 weeks) of 6,025 unemployment insurance benefit recipients in New Jersey, conducted between 2009 and 2010.[44] Importantly, the survey contains data on job search behavior, both in the form of weekly recall questions about activity over the entire previous week, as well as a once-weekly time diary. Most relevantly for studying whether the internet makes search less lumpy, the SUWNJ contains both a question asking respondents whether they used the internet to search for work in the past week and information on whether or not a respondent searched on the day prior to the interview via the time diaries.

The motivation for the empirical strategy and specification below is straightforward: If online job search is smoother than traditional job search for reasons described above, then the probability that a respondent searches on any given day should be higher for a person who uses the internet for job search than for a person who does not, all else equal. Following this logic, let s_{it}^{TD} denote the time-diary measure of total search time by respondent i on the day prior to the week-t interview, s_{it}^{WR} the weekly-recall measure of total search time over the past week, OJS_{it} a dummy for whether a respondent reports having used the internet for job search in the past week, d_{it} unemployment duration, τ_t a calendar-week fixed effect, and η_i a respondent fixed effect. I then estimate variations on a linear probability model of the form:

$$Pr(s_{it}^{\mathrm{TD}} > 0) = \alpha + \beta \mathrm{OJS}_{it} + \gamma s_{it}^{\mathrm{WR}} + \zeta d_{it} + \tau_t + \eta_i + \epsilon_{it}. \tag{56}$$

The coefficient β on OJS_{it} is the principal object of interest. In all specifications, I restrict attention to respondents between the ages of 20 and 60 who have not accepted a job offer. Moreover, I follow Krueger and Mueller (2011) and drop time diaries in which fewer than four distinct activities are reported or in which three or more hour-long episodes are left incomplete. All regressions also include day-of-week fixed effects.

Table 4 reports the main results. The first column ("Spec. 1") corresponds to the baseline regression described above; the second column ("Spec. 2") adds a control for the total number of applications that a respondent reports submitting during the week prior to the survey; and the third column

[43] See the Introduction for examples of technologies that reduce fixed costs and enable monitoring.

[44] Complete survey data and documentation can be obtained from https://dss.princeton.edu/catalog/resource1350.

("Spec. 3") adds a set of 12 dummies for whether or not a respondent used various specific search methods during the week prior to the survey.[45] The second two columns are important because online job search is different from more traditional methods of search in ways other than those described above that could potentially also affect the likelihood that a respondent searches on a random day.[46] This makes controlling for other aspects of a respondent's search behavior, such as the number of applications they submit ("Spec. 2") or the methods they use ("Spec. 3"), important for identifying the search-smoothing effect described above.

Table 4: Smooth search

	Spec. 1	Spec. 2	Spec. 3
Online search (β)	0.16***	0.14***	0.11***
	(0.03)	(0.03)	(0.03)
Total search (γ)	0.03***	0.02***	0.01*
	(0.00)	(0.00)	(0.00)
Unempl. duration (ζ)	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)
Total app'ns		0.05***	0.04***
		(0.01)	(0.01)
Fixed effects:			
Respondent:	×	×	×
Calendar time:	×	×	×
Day of week:	×	×	×
Search methods:			×
Observations	17,936	17,028	17,028

Robust standard errors in parentheses.

Notes: Standard errors clustered at the individual level. Sample restricted to respondents who have never accepted a job offer and who are between the ages of 20 and 60. All specifications use survey weights.

The results indicate that engaging in online search increases the probability of searching on the previous day by at least 11 percentage points, an effect that is highly significant across specifications. This is consistent with online search allowing job seekers to smooth out their search across days rather than devoting, e.g., a single day of the week to search to economize on fixed costs. Moreover, the results are robust to various alternative specifications, sample selection criteria, and inclusion of additional controls. For example, controlling for reported reservation wages, excluding long-term unemployed workers, and controlling for the amount of time that a respondent reports spending on each of the 12 search methods in the past week (rather than simply the *total* amount of time and indicators for whether they engaged in each of the specific activities) do not affect the results.

[45] The 12 options are contacting an employer directly, contacting a public employment agency, contacting a private employment agency, contacting friends or relatives, contacting a school/university employment center, checking union/professional registers, attending a job training program/course, placing or answering ads, going to an interview, sending out resumes/completing applications, looking at ads, and a category for all other methods.

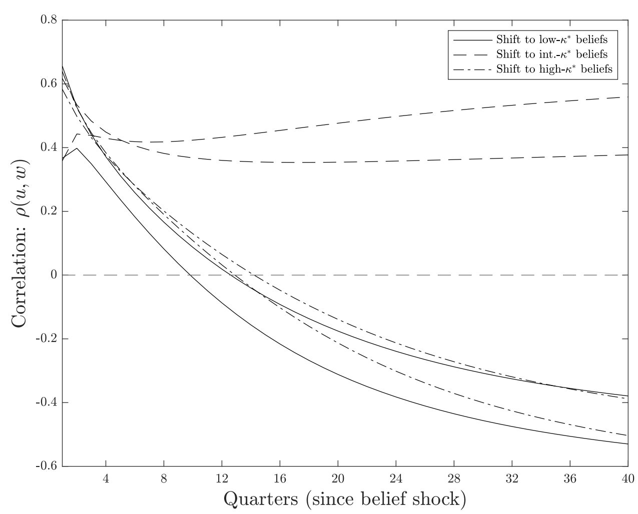
[46] See the Introduction for a partial list of literature exploring other effects of online search.

6.2 Macro evidence: Labor supply shocks are increasingly important

Macroeconomists have long been interested in identifying the contribution of labor supply shocks to aggregate fluctuations. This task is approached in the SVAR literature by proposing theoretically motivated restrictions that enable labor supply shocks to be plausibly disentangled from other sources of aggregate fluctuations.[47] As discussed in Section 5, the model described in this paper can be understood as a theory of belief-driven labor supply shocks. Given that monitoring has presumably become increasingly important since the popularization of online job search around the turn of the century, it is natural to ask whether labor supply shocks have come to contribute more to aggregate fluctuations in the past 20 years than they did previously.[48]

Before proceeding, it is first necessary to return briefly to the analysis of Section 5.3 (where I considered the possibility of unanticipated shocks to beliefs shifting the economy between equilibria) to understand how belief shocks can be identified in aggregate data. To this end, Figure 9 plots the correlation between wages and unemployment extending out to various horizons following each of the six possible shocks to beliefs considered in Section 5.3.[49] Solid and dotted-dashed lines correspond to the shocks that shift the economy to the low-monitoring and high-monitoring trajectories, respectively; dashed lines correspond to the shocks that shift the economy to the intermediate-monitoring trajectory. For the purposes of the analysis in this section, the critical

Figure 9: Identification of belief shocks



Notes: The figure depicts the correlations between unemployment and wages computed over various horizons (1-40) following the six belief shocks. For all six shocks, at all horizons less than two years (eight quarters), the implied correlation between wages and unemployment following the shock is positive, suggesting that short-run sign restrictions will be effective in identifying belief shocks and that such shocks will be identified as labor supply shocks.

[47] See, for example, Blanchard and Diamond (1989), Chang and Schorfheide (2003), Peersman and Straub (2009), and Foroni et al. (2018).

[48] This possibility is consistent with the results in Foroni et al. (2018), who observe that the importance of labor supply shocks is magnified when their sample is extended to include more recent data.

[49] For example, the correlations corresponding to quarter 12 represent the correlations between unemployment and wages over the length of the first 12 quarters (36 months) following the shock.

observation is that all six of the shocks induce a positive correlation between unemployment and wages (and thus a negative correlation between output and wages) at all horizons of less than two years, consistent with the interpretation as labor supply shocks. After two years, the story is more complicated: Shocks shifting the economy to the intermediate-monitoring trajectory continue to induce a positive correlation, whereas the other shocks induce a negative correlation.[50] This suggests that short-run restrictions on the relationship between unemployment (or output) and wages will be likely to identify belief shocks resulting from monitoring as labor supply shocks, whereas long-run restrictions may struggle to do so.

Accordingly, I estimate a standard three-variable VAR using a short-run sign restriction approach similar to that in Peersman and Straub (2009) and further refined by Foroni et al. (2018). Specifically, following Foroni et al. (2018), I estimate a VAR with output (quarterly real output in the non-farm business sector from the BLS), the price level (GDP deflator), and a measure of real wages (total private average hourly earnings of production and non-supervisory employees, deflated by the GDP deflator), each measured at a quarterly frequency. The VAR includes four lags and is estimated in (log) levels. Following Foroni et al. (2018), I assume that (i) aggregate demand shocks move output and the price level in the same direction on impact, (ii) technology shocks move output and the price level in opposite directions on impact (thus identifying them from aggregate demand shocks), and (iii) technology shocks move the real wage in the opposite direction as labor supply shocks on impact.[51]

Reflecting the role of the internet in enabling job seekers to monitor vacancies, I split the post-1985 sample into two non-overlapping periods: 1985Q1-1999Q4 and 2001Q1-2019Q4.[52] I then estimate the VAR separately on each sample and report the implied forecast error variance decomposition. Figure 10 reports the results for output and wages at horizons of up to 10 years (40 quarters).

Panel 10a reports the contribution of labor supply shocks to variation in output. Panel 10b reports the contribution of labor supply shocks to variation in the real wage. In both panels, the dark shaded area corresponds to the pre-2000 sample, while the sum of the dark and light shaded areas corresponds to the post-2000 sample. Thus, in both cases, the contribution of labor supply shocks is markedly larger over the past 20 years than it was over the preceding 15, prior to the advent of online job search. Furthermore, this observation holds at all horizons, notwithstanding the conventional wisdom that labor supply shocks are more relevant at long horizons. The necessity of splitting the sample around the time at which online search technologies began to become widespread (i.e., 2000) in order to separately analyze the pre- and post-online job search economies makes the particular quantitative features of the results in Figure 10 somewhat sensitive to details of the VAR specification. However, the qualitative observation that labor supply shocks seem to play an increased role in explaining the dynamics of output and wages since 2000 is a robust feature of the data.

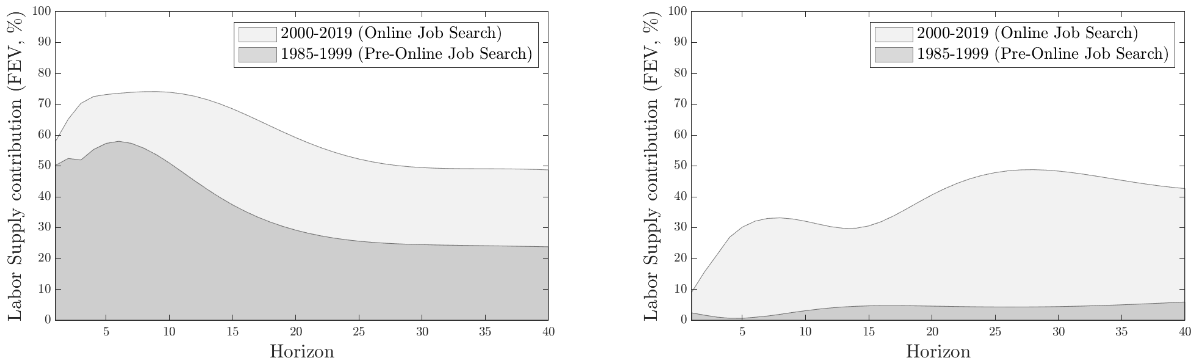
While not conclusive, this evidence is consistent with the idea put forward in this paper that online job search technologies—and monitoring technologies in particular—can expose the economy to

[50] I discuss the reason for this in Section 5.3.

[51] Despite the fact that unemployment (and thus output) is predetermined in the model, implying that it does not move on impact following a shock, because the period length in the model is one month while the period length in the data is one quarter, quarterly impact restrictions will be sufficient to identify belief-induced labor supply shocks.

[52] The use of a post-1985 sample is intended to focus on the "Great Moderation" period. This is also the sample used by Foroni et al. (2018) and similar to the years I use for calibration in Section 4.

Figure 10: Variance decomposition: Pre- and post-online job search



(a) Output (b) Wage

Notes: Panel (a) depicts the contribution of labor supply shocks to variation in output at various horizons. Panel (b) depicts the contribution of labor supply shocks to variation in wages at various horizons. In both panels, the dark area corresponds to the pre-2000 sample, while the total area (light and dark grey combined) corresponds to the post-2000 sample.

belief-driven fluctuations in labor supply that would otherwise not exist.

7 Conclusion

This paper conceptualizes search as a monitoring decision and studies the implications for equilibrium labor market dynamics. I first show that monitoring leads to a novel source of strategic complementarities in search and thus multiple equilibria. The core intuition is that monitoring technologies enable workers who actively monitor vacancies to find and apply for jobs before workers who do not, leading to a rat race among unemployed workers trying to avoid being last in line for work. I then embed this mechanism in a quantitative model of the labor market and show that, when the job creation process is not too elastic, the model has multiple steady-state equilibria and also multiple dynamic equilibria in the sense that, from a range of initial positions, the economy can converge to different steady states. This implies that endogenous changes in workers' beliefs can permanently alter the trajectory of the economy. Moreover, the model offers a parsimonious account of several important features of the recovery from the Great Recession that are difficult to reconcile with traditional models.

References

Acharya, S., J. Bengui, K. Dogra, and S. L. Wee (2018). Slow recoveries and unemployment traps: Monetary policy in a time of hysteresis. *Working Paper*.

Albrecht, J., B. Decreuse, and S. Vroman (2017). Directed search with phantom vacancies. *Unpubished manuscript*.

Atolia, M. and E. Buffie (2009). Reverse shooting made easy: Automating the search for the global nonlinear saddle path. *Manuscript*. Manuscript.

Autor, D. (2001). Wiring the labor market. *Journal of Economic Perspectives*.

Barron, J., M. Berger, and D. Black (1997). Employer search, training, and vacancy duration. *Economic Inquiry*.

Beaudry, P., D. Green, and B. Sand (2013). How elastic is the job creation curve? *Unpublished manuscript*.

Beaudry, P., D. Green, and B. a. Sand (2018). In search of labor demand. *American Economic Review*.

Belot, M., P. Kircher, and P. Muller (2019). Providing advice to job seekers at low cost: An experimental study on online advice. *Review of Economic Studies*.

Blanchard, O. and P. Diamond (1989). The beveridge curve. *Brookings Papers on Economic Activity*.

Brunner, M. and H. Strulik (2002). Solution of perfect foresight saddle point problems: a simple method and applications. *Journal of Economic Dynamics and Control*.

Cairo, I., S. Fujita, and C. Morales-Jimenez (2021). The cyclicality of labor force participation flows: The role of labor supply elasticities and wage rigidity. *Review of Economic Dynamics*.

Carnevale, A., T. Jayasundera, and D. Repnikov (2014). Understanding online job ads data. Technical report, Georgetown University.

Chang, Y. and F. Schorfheide (2003). Labor-supply shifts and economic fluctuations. *Journal of Monetary Economics*.

Cheron, A. and B. Decreuse (2017). Matching with phantoms. *Review of Economic Studies*.

Chodorow-Reich, G. and L. Karabarbounis (2016). The cyclicality of the opportunity cost of employment. *Journal of Political Economy*.

Coles, M. and A. Moghaddasi Kelishomi (2018). Do job destruction shocks matter in the theory of unemployment? *American Economic Journal: Macroeconomics*.

Coles, M. and E. Smith (1998). Marketplaces and matching. *International Economic Review*.

Davis, S. and B. Samaniego de la Parra (2020). Application flows. *Working paper*.

Diamond, P. (1982a). Aggregate demand management in search equilibrium. *Journal of Political Economy 90*(5).

Diamond, P. (1982b). Wage determination and efficiency in search equilibrium. *Review of Economic Studies*.

Dube, A. (2019). Impacts of minimum wages: Review of the international evidence. Technical report, HM Treasury.

Eeckhout, J. and I. Lindenlaub (2018). Unemployment cycles. *American Economic Journal: Macroeconomics*.

Faberman, J. and M. Kudlyak (2016). What does online job search tell us about the labor market? *Federal Reserve Bank of Chicago Economic Perspectives*.

Faberman, J. and M. Kudlyak (2019). The intensity of job search and search duration. *American Economic Journal: Macroeconomics*.

Fonseca, R., P. Lopez-Garcia, and C. Pissarides (2000). Entrepreneurship, start-up costs and employment. *European Economic Review*.

Foroni, C., F. Furlanetto, and A. Lepetit (2018). Labor supply factors and economic fluctuations. *International Economic Review*.

Freeman, R. (2002). The labour market in the new information economy. *Oxford Review of Economic Policy*.

Fujita, S. and G. Ramey (2005). The dynamic beveridge curve. *Federal Reserve Bank of Philadelphia Working Papers*.

Fujita, S. and G. Ramey (2012). Exogenous versus endogenous separation. *American Economic Journal: Macroeconomics*.

Hall, R. and P. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American Economic Review*.

Hamermesh, D. (1993). *Labor Demand*. Princeton University Press.

Howitt, P. and P. McAfee (1992). Animal spirits. *American Economic Review*.

Judd, K. (1999). *Numerical Methods in Economics*. MIT Press.

Kaplan, G. and G. Menzio (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy*.

Kroft, K. and D. Pope (2014). Does online search crowd out traditional search and improve matching efficiency? evidence from craigslist. *Journal of Labor Economics*.

Krueger, A. and A. Mueller (2011). Job search, emotional well-being and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data. *Brookings Papers on Economic Activity 42*, 1–81.

Krusell, P., T. Mukoyama, R. Rogerson, and A. Sahin (2017). Gross worker flows over the business cycle. *American Economic Review*.

Kuhn, P. (2003). The internet and matching in labor markets. In *New Economy Handbook*.

Kuhn, P. and H. Mansour (2013). Is internet job search still ineffective? *The Economic Journal*.

Kuhn, P. and M. Skuterud (2004). Internet job search and unemployment durations. *American Economic Review*.

Landais, C., P. Michaillat, and E. Saez (2018). A macroeconomic approach to optimal unemployment insurance: Theory. *American Economic Journal: Economic Policy*.

Menzio, G. and P. Martellini (2020). Declining search frictions, unemployment and growth. *Journal of Political Economy 128*.

Mortensen, D. (1982). The matching process as a noncooperative bargaining game. In U. of Chicago Press (Ed.), *The Economics of Information and Uncertainty*.

Mukoyama, T., C. Patterson, and A. Sahin (2018). Job search behavior over the business cycle. *American Economic Journal: Macroeconomics*.

Peersman, G. and R. Straub (2009). Technology shocks and robust sign restrictions in a euro area svar. *International Economic Review*.

Petrongolo, B. and C. Pissarides (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*.

Pissarides, C. (1985). Short-run dynamics of unemployment, vacancies, and real wages. *American Economic Review*.

Pissarides, C. (1992). Loss of skill during unemployment and the persistence of employment shocks. *Quarterly Journal of Economics*.

Pries, M. and R. Rogerson (2022). Declining worker turnover: The role of short-duration employment spells. *American Economic Journal: Macroeconomics*.

Schaal, E. and M. Taschereau-Dumouchel (2016). Aggregate demand and the dynamics of unemployment. *Working Paper*.

Shapiro, M. and M. Watson (1988). Sources of business cycle fluctuations. *NBER Macroeconomics Annual*.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*.

Shimer, R. (2012). Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*.

Sterk, V. (2016). The dark corners of the labor market. *Discussion Paper*.

Stevenson, B. (2009). The internet and job search. In *Labor Market Intermediation*. University of Chicago Press.

Toohey, D. (2021). Work search requirements in general equilibrium. *Working Paper*.

Appendices

A Monitoring game

A.1 Proposition 1: Sufficient conditions for multiplicity and uniqueness.

Proof. I first establish the sufficient condition for multiplicity and then establish the sufficient condition for uniqueness.

Sufficient condition for multiplicity: The proof proceeds by first showing that for any $\theta > 1/2$ there are multiple equilibria provided search costs are bounded between $[(1-\theta)(w-b), \theta(w-b)]$, from which it follows immediately from the assumption that $\bar{\kappa} < w-b$ that there must exist some $\bar{\theta}$ such that for any $\theta \in (\bar{\theta}, 1)$, there is multiplicity. The interval $[(1-\theta)(w-b), \theta(w-b)]$ is non-empty only if $\theta \geq \frac{1}{2}$. Suppose $\kappa^* = \underline{\kappa}$ so that $F(\kappa^*) = 0$ and no workers search. Then, using (10) and the fact that $F(\kappa^*) = 0 < \theta$, if $\underline{\kappa} > (1-\theta)(w-b)$ then $\kappa_i^*(\kappa^*) = \kappa^L(\kappa^*) = \underline{\kappa}$, so $\kappa^* = \underline{\kappa}$ is an equilibrium. Suppose instead that $\kappa^* = \bar{\kappa}$ so that $F(\kappa^*) = 1$ and all workers search. Then, again using (10) and the fact that $F(\kappa^*) = 1 > \theta$, if $\bar{\kappa} < \theta(w-b)$ then $\kappa_i^*(\kappa^*) = \kappa^H(\kappa^*) = \bar{\kappa}$, so $\kappa^* = \bar{\kappa}$ is an equilibrium. Thus, there are multiple equilibria if $\underline{\kappa} > (1-\theta)(w-b)$ and $\bar{\kappa} < \theta(w-b)$, i.e., if search costs are bounded between $[(1-\theta)(w-b), \theta(w-b)]$.

Sufficient condition for uniqueness: The proof proceeds by showing that there exists a value $\underline{\theta}$ such that for any $\theta \in (0, \underline{\theta})$ there must exist a unique equilibrium. To establish existence, first note that (12) implies that there must exist some $\underline{\theta}_0 > 0$ such that for any $\theta < \underline{\theta}_0$, $\kappa^U(\kappa^*) = \frac{\theta}{F(\kappa^*)}(w-b)$ for all $\kappa^* \in [\kappa^c, \bar{\kappa}]$. Thus, because κ^U is weakly decreasing and continuous on $[\kappa^c, \bar{\kappa}]$ for any such θ, it is sufficient to establish that (i) $\frac{\theta(w-b)}{F(\bar{\kappa})} \leq \bar{\kappa}$ and (ii) $\frac{\theta(w-b)}{F(\kappa^c)} \geq \kappa^c$ (i.e., that the best-response function crosses the 45-degree line). Because $F(\bar{\kappa}) = 1$, (i) is true for any $\theta < \underline{\theta}_1 \equiv \frac{\bar{\kappa}}{w-b} \in (0,1)$, where $\frac{\bar{\kappa}}{w-b} \in (0,1)$ follows from the assumptions that $\bar{\kappa} < w-b$ and that both $w-b$ and $\bar{\kappa}$ are strictly positive. Because $\kappa^c \equiv F^{-1}(\theta)$, (ii) is true if $w-b > F^{-1}(\theta)$. Because F must be increasing and by assumption $w-b > \bar{\kappa}$, this holds for any $\theta < 1$. Thus, equilibrium exists for $\theta < \min\{\underline{\theta}_0, \underline{\theta}_1\} > 0$. To establish uniqueness, note that in the limit as $\theta \to 0$, the only possible equilibrium such that $\theta \geq F(\kappa^*)$ corresponds to $\kappa^* = \underline{\kappa}$, which, by (11), requires $w-b < \underline{\kappa} < \bar{\kappa}$, a contradiction. Thus, there exists some $\underline{\theta} < \min\{\underline{\theta}_0, \underline{\theta}_1\}$ such that there is a unique equilibrium for any $\theta < \underline{\theta}$. Furthermore, that equilibrium must be such that $\theta < F(\kappa^*)$. □

A.2 Proposition 2: Impossibility of multiplicity in traditional model

Proof. From the best-response function in (3), a necessary condition for multiplicity is that $p^h - p^l$ is increasing in κ^* (that is, increasing returns are necessary for multiplicity). Because $z < 1$, (13) implies that $\tilde{s}(\kappa^*)$ is increasing in κ^*, while (15) and (16) imply that $p^h - p^l$ is decreasing in $\tilde{s}(\kappa^*)$. Thus, $p^h - p^l$ is decreasing in κ^*, so there cannot be multiple equilibria. □

A.3 Special case: Uniform κ_i

Consider the special case in which the value of leisure/home production κ_i is Uniformly distributed on $[0, \bar{\kappa}]$ so that $F(\kappa_i) = \frac{\kappa_i}{\bar{\kappa}}$. This special case allows for the Nash equilibria of the game to be solved in closed form, which makes it easy to study the model's comparative statics across equilibria.

Specifically, using (10), it is straightforward to see that there are potentially three equilibria,

$$\kappa^*_{\text{Low}} = \frac{\bar{\kappa}}{2} - \sqrt{\left(\frac{\bar{\kappa}}{2}\right)^2 - (1-\theta)(w-b)\bar{\kappa}} \tag{A.1}$$

$$\kappa^*_{\text{Int.}} = \frac{\bar{\kappa}}{2} + \sqrt{\left(\frac{\bar{\kappa}}{2}\right)^2 - (1-\theta)(w-b)\bar{\kappa}} \tag{A.2}$$

$$\kappa^*_{\text{High}} = \begin{cases} \sqrt{\theta(w-b)\bar{\kappa}} & \text{if} \quad \bar{\kappa} > \theta(w-b) \\ \bar{\kappa} & \text{if} \quad \bar{\kappa} \leq \theta(w-b) \end{cases} \tag{A.3}$$

where the first two potential equilibria (κ^*_{Low} and $\kappa^*_{\text{Int.}}$) correspond to a situation in which workers are the short end of the monitoring phase ($F(\kappa^*) \leq \theta$) and the third potential equilibrium (κ^*_{High}) corresponds to a situation in which vacancies are the short end of the monitoring phase ($F(\kappa^*) > \theta$).

Inspection of (A.1) and (A.2) reveals several interesting features of the game. First, consistent with Proposition 1, multiplicity is guaranteed as $\theta \to 1$ and precluded for any $\theta < 0.5$. Second, conditional on all three equilibria existing, a higher employment rent $(w - b)$ will tend to increase monitoring in the high- and low-monitoring equilibria, but decrease monitoring in the intermediate-monitoring equilibrium. Third, a sufficiently low employment rent will tend to eliminate multiplicity and render κ^*_{Low} the unique equilibrium.

B Traditional model

This appendix describes a traditional model of matching reflecting how search is typically modeled in the literature—as a decision concerning the intensity of effort rather than as a decision concerning whether to monitor new jobs.

B.1 Search as variable effort

The traditional approach to modeling job search decisions in quantitative macroeconomic models is to assume that searchers and non-searchers are perfect substitutes in the match function. In this case, non-searchers are simply ineffective searchers who match at a rate proportional to that of searchers. Recent examples of such models include Krusell et al. (2017) and Cairo et al. (2021). The description of search and matching below reflects this interpretation.

Consider a model in which workers choosing to search exert one unit of effort, whereas workers choosing not to search exert $z < 1$ units of effort. "Effort" here can be interpreted as the probability of participating in the market. Taking this interpretation and dropping time subscripts for ease of notation, the effective measure of searchers in the economy, s, can be defined as

$$s = F(\kappa^*)u + (1 - F(\kappa^*))zu \tag{B.1}$$

$$= \Big[F(\kappa^*) + (1 - F(\kappa^*))z\Big]u \tag{B.2}$$

$$= \tilde{s}u. \tag{B.3}$$

where $\tilde{s}$ can be interpreted as average search effort in the economy.

B.2 Monitoring Game

The traditional characterization of search described above can be immediately embedded in the basic environment described in Section 2. Specifically, the search decision described in Section 2.2.1 does not depend on interpreting search as monitoring, and thus continues to hold when search is interpreted instead as a choice of intensity. The only change, then, are the match rates in equations (7) and (8). In Section 2, I discuss how the changes implied by this more traditional model of search affect the existence of increasing returns and thus multiple equilibria.

B.3 Dynamic model

Adopting the functional form of the match function from Section 3, the total number of matches in this economy is given by

$$m_t = \mu(v_t, s_t) = v_t \left[1 - e^{-\psi \frac{s_t}{v_t}}\right]. \tag{B.4}$$

The implied match rates for searchers (high-effort) and non-searchers (low-effort), respectively, are thus

$$p_t^h = \frac{v_t}{s_t}\left[1 - e^{-\psi \frac{s_t}{v_t}}\right] \tag{B.5}$$

$$p_t^l = z\frac{v_t}{s_t}\left[1 - e^{-\psi \frac{s_t}{v_t}}\right]. \tag{B.6}$$

The average match rate for workers, in turn, is

$$p_t = F(\kappa_t^*)p_t^h + (1 - F(\kappa_t^*))p_t^l \tag{B.7}$$

$$= \theta_t \left[1 - e^{-\psi \frac{\tilde{s}_t}{\theta_t}}\right] \tag{B.8}$$

where the second line uses the expression for s_t in (B.3), (B.5), (B.6), and $\theta_t \equiv v_t/u_t$. Thus, changes in average search effort are isomorphic to changes in ψ, the parameter governing frictions in the matching process. Note that this implies that it is possible to write the difference between the match rate associated with searching and the match rate associated with not searching (which governs the search decision) as

$$p_t^h - p_t^l = \frac{1-z}{\tilde{s}_t}p_t. \tag{B.9}$$

Finally, the average match rate for firms in the traditional model is just

$$q_t = p_t/\theta_t = 1 - e^{-\psi \frac{\tilde{s}_t}{\theta_t}}. \tag{B.10}$$

The foregoing can be used to calibrate the model using the exact same moments that are used to calibrate the model in Section 4 as described in Appendix C.

C Calibration

This appendix provides further details on the calibration of ψ, δ, ν, $\tilde{\eta}$, μ_κ, and σ_κ. All other parameters are calibrated directly to match values in the data and existing literature as described in Section 4. The six moments used to pin down these remaining six parameters are: $\bar{d}$ (average vacancy duration), $\bar{p}$ (average match rate of workers), $\bar{u}$ (unemployment rate), $\bar{z} \equiv bw + \int_{\kappa^*}^{\infty} \kappa dF(\kappa)$ (average flow value of unemployment), $\bar{F}$ (share of workers monitoring), and $\bar{\epsilon}_{e,w} \equiv \frac{d\ln(e)}{d\ln(w)}$ (elasticity of employment with respect to the wage).

As described in Section 4, I calibrate the model in two steps. In the first step, I consider the model without monitoring (which is just a standard DMP-style model with a generalized job-creation process), and calibrate the parameters to match moments from macroeconomic data from 1985-1999 and existing studies based primarily on data from before 2000.[53] This leaves only the standard deviation of the leisure/home production distribution to be calibrated. In the second step, I consider the model with monitoring, and calibrate this remaining parameter to match a back-of-envelope calculation of the share of monitoring workers in the early 2000s based on data from Stevenson (2009).[54]

Step #1: Pre-2000 model (no monitoring)

I first calibrate all of the standard parameters to a version of the model without monitoring, in which I replace the optimality condition for κ^* in (34) with $\kappa^* = 0$.

Match efficiency (ψ). In the absence of monitoring, workers only match in the aftermarket at the end of each period. This implies $\mathring{m} = 0$, from which it follows that the total number of matches is given by $m = \mathring{m} + \hat{m} = \hat{m}$. Thus, using equation (46),

$$m = (v - F(\kappa^*)u)\left[1 - e^{-\psi \frac{u - F(\kappa^*)u}{v - F(\kappa^*)u}}\right] \tag{C.1}$$

$$= u(\theta - F(\kappa^*))\left[1 - e^{-\psi \frac{1 - F(\kappa^*)}{\theta - F(\kappa^*)}}\right] \tag{C.2}$$

$$= u\theta\left[1 - e^{-\frac{\psi}{\theta}}\right] \tag{C.3}$$

where the last line follows from the fact that $F(\kappa^*) = 0$ if no workers monitor. Thus, the average match rate for vacancies, q, is

$$q \equiv m/v \tag{C.4}$$

$$= 1 - e^{-\frac{\psi}{\theta}}. \tag{C.5}$$

Average vacancy duration (in days), $\bar{d}$, pins down the monthly average match rate for firms via

$$q = 1 - \left(1 - \frac{1}{\bar{d}}\right)^{31+1} \tag{C.6}$$

[53] The idea here is to capture the portion of the Great Moderation during which monitoring technologies were not in wide use. See the main text for further discussion.

[54] A small technicality should be pointed out before proceeding: Calibrating the model without monitoring requires a value for the mean of the leisure/home production distribution, as this affects the bargained wage. Because I assume this distribution to be Log-Normal, it has two parameters, so the model without monitoring is consistent with a continuum of combinations of the two parameters μ_κ and σ_κ that yield the targeted mean. In order to pin down the second parameter of this distribution, I consider the model with monitoring, and choose the second parameter of this distribution to match my calculation of the share of monitoring workers as described in the text.

which, in turn, can be used to solve for θ using the average match rate for workers, $\bar{p}$:

$$\theta = \bar{p}/q. \tag{C.7}$$

Using (C.6) and (C.7), equation (C.5) can be solved for the match efficiency parameter, ψ:

$$\psi = -\theta \ln(1-q). \tag{C.8}$$

Separation rate (δ). Given $\bar{p}$ and $\bar{u}$, equation (19) can be solved for the separation rate, δ:

$$\delta = \frac{\bar{p}}{\bar{p} + 1/\bar{u} - 1}. \tag{C.9}$$

Job creation ($\tilde{\eta}$, ν). I next calibrate the parameter scaling the vacancy-creation process ($\tilde{\eta}$) and the parameter governing the elasticity of the vacancy-creation process (ν).

Entry depends on the steady-state value of a new vacancy. The value of a new vacancy, in turn, depends on the match rate for new vacancies, q^n, the match rate for old vacancies, q^o, and the value of the wage, w. I solve for these objects in turn.

Using $\theta \equiv v/\bar{u} = \bar{p}/q$, it is possible to pin down the steady-state level of vacancies,

$$v = \frac{\bar{p}}{q}\bar{u}. \tag{C.10}$$

Given v and q, the law of motion for vacancies in (20) pins down the number of new vacancies via

$$v^n = v(1-(1-\delta)(1-q)). \tag{C.11}$$

From here, it is possible to solve for all of the relevant match rates. Specifically, since the model is being calibrated to the pre-2000 period in which there was no monitoring technology, it must be that $\mathring{m} = 0$. Furthermore, given $\bar{u}$ and the value of θ solved for above, (C.3) pins down $\hat{m}$ and m (the number of matches in the aftermarket and the total number of matches, which are equal). Using these and equations (23) and (26), the steady-state match rates for vacancies in the monitoring phase and aftermarket, respectively, are

$$\mathring{q} = \mathring{m}/v^n = 0 \tag{C.12}$$
$$\hat{q} = \hat{m}/(v - \mathring{m}) = q. \tag{C.13}$$

From (41) and (42), the steady-state match rates for new and old vacancies, respectively, are

$$q^n = \mathring{q} + (1-\mathring{q})\hat{q} = q \tag{C.14}$$
$$q^o = \hat{q} = q. \tag{C.15}$$

Likewise, from equations (22) and (25), the steady-state match rates for workers in the monitoring phase and aftermarket, respectively, are

$$\mathring{p} = \mathring{m}/u^m = 0 \tag{C.16}$$

$$\hat{p} = \hat{m}/(u - \mathring{m}) = \bar{p}. \tag{C.17}$$

From (28) and (29), the steady-state match rates for workers who choose to monitor and who choose not to monitor, respectively, are

$$p^m = \mathring{p} + (1 - \mathring{p})\hat{p} = \bar{p} \tag{C.18}$$

$$p^w = \hat{p} = \bar{p}. \tag{C.19}$$

In short, without monitoring, there is only one match rate for workers, $\bar{p}$, and one match rate for vacancies, q, and the model reduces to a standard DMP model. The next step is to pin down the steady-state wage, w. To do so, observe that in general (that is, with or without monitoring), the average flow value of non-employment is

$$E[z] \equiv bw + \int_{\kappa^*}^{\infty} \kappa dF(\kappa) \tag{C.20}$$

where the first term is the UI payment and the second term is the average value of leisure or home production across all non-employed workers. The partial expectation reflects that monitoring entails forgoing leisure or home production. Thus, in the absence of monitoring, $\kappa^* = 0$, which implies $\int_{\kappa^*}^{\infty} \kappa dF(\kappa) = E[\kappa]$. Using (43) and (44), it is possible to solve for the Nash bargained wage as a function of steady-state objects that have already been solved for and the partial expectation in (C.20):

$$w = \frac{\left(\frac{\chi}{1-\beta(1-\delta)(1-q^o)}\right)(y+c) + \left(\frac{1-\chi}{1-\beta(1-\delta)(1-p)}\right)\int_{\kappa^*}^{\infty} \kappa dF(\kappa)}{\left(\frac{1-(1-b)u}{1-u}\right)\left(\frac{\chi}{1-\beta(1-\delta)(1-q^o)}\right) + (1-b)\left(\frac{1-\chi}{1-\beta(1-\delta)(1-p)}\right)}. \tag{C.21}$$

From here, (C.20) can be used to eliminate the partial expectation from (C.21) and solve for the steady-state wage as a function of b and $\overline{E[z]}$, both of which are taken from the data:

$$w = \frac{\left(\frac{\chi}{1-\beta(1-\delta)(1-q^o)}\right)(y+c) + \left(\frac{1-\chi}{1-\beta(1-\delta)(1-\bar{p})}\right)\overline{E[z]}}{\left(1 + b\left(\frac{\frac{1-\chi}{1-\beta(1-\delta)(1-\bar{p})}}{\left(\frac{1-(1-b)\bar{u}}{1-\bar{u}}\right)\frac{\chi}{1-\beta(1-\delta)(1-q^o)} + (1-b)\frac{1-\chi}{1-\beta(1-\delta)(1-\bar{p})}}\right)\right)\left[\left(\frac{1-(1-b)\bar{u}}{1-\bar{u}}\right)\frac{\chi}{1-\beta(1-\delta)(1-q^o)} + (1-b)\left(\frac{1-\chi}{1-\beta(1-\delta)(1-\bar{p})}\right)\right]}. \tag{C.22}$$

As described in Section 4, I use the value of $\overline{E[z]}$ computed by Hall and Milgrom (2008), which is also in the middle of the range of values of the opportunity cost of employment identified by Chodorow-Reich and Karabarbounis (2016). With a value for w, it is possible to solve for the steady-state tax, τ, using (44),

$$\tau = \frac{bw\bar{u}}{1-\bar{u}}. \tag{C.23}$$

Next, solving for the value of a filled job using (40),

$$J = \frac{y - w - \tau}{1 - \beta(1-\delta)} \tag{C.24}$$

which allows us to solve for the value of an old vacancy and the value of a new vacancy, respectively:

$$V^o = \frac{-c + \beta(1-\delta)q^o J}{1 - \beta(1-\delta)(1-q^o)} \tag{C.25}$$

$$V^n = -c + \beta(1-\delta)\left[q^n J + (1-q^n)V^o\right]. \tag{C.26}$$

Finally, I use the vacancy-creation condition in (37) to calibrate the two parameters governing the vacancy-creation process,

$$v^n = \tilde{\eta}(V^n)^\nu \tag{C.27}$$

where $\tilde{\eta} \equiv \eta\bar{\xi}^{-\nu}$.[55] I calibrate ν to match estimates of the wage elasticity of employment. Specifically, I numerically solve for the value of ν that yields a steady-state elasticity, $\frac{d\ln e}{d\ln w}$, equal to -0.3, a value in the middle of the range of estimates in the literature:

$$\nu^* = \left\{\nu \middle| \bar{\epsilon}_{e,w} = \frac{d\ln e}{d\ln w}(\nu)\right\} \tag{C.28}$$

where the elasticity on the right-hand side of the restriction, $\frac{d\ln e}{d\ln w}$, is expressed as an explicit function of ν to stress that ν is critical for the model-implied value of this elasticity. Finally, I calibrate $\tilde{\eta}$ by solving (C.27) for $\tilde{\eta}$ given the value of ν obtained above, yielding

$$\tilde{\eta} = v^n/(V^n)^\nu. \tag{C.29}$$

Step #2: Post-2000 model (monitoring)

I next calibrate the parameters of the leisure/home production distribution, μ_κ and σ_κ.[56]To do so, I proceed in two sub-steps:

2a) Replace the assumption in Step #1 that $\kappa^* = 0$ with an assumption that κ^* is instead governed by the optimality condition in (34).

2b) Under the parameter values identified in Step #1 above, solve for the steady state of the model with monitoring, using two restrictions to pin down μ_κ and σ_κ:

   (i) $bw + \int_0^\infty \kappa dF(\kappa) = \bar{E[z]}$ [as in Step #1]

   (ii) $F(\kappa^*) = \bar{F}$

   If there are multiple steady states, select the one with the lowest level of monitoring.

Restriction (i) requires that, when no workers monitor, μ_κ and σ_κ must be such that the steady-state flow value of unemployment is equal to the value from the data, $\bar{E[z]}$. This implies that the monitoring model is consistent with the no-monitoring model in the sense that, when $\kappa^* = 0$, a steady state identical to the one obtained in Step #1 above is recovered. Restriction (ii) requires that the share of monitoring workers in the low-monitoring steady-state be equal to the value

[55] See Footnote 26 for an explanation of why I calibrate $\tilde{\eta}$ instead of η.

[56] As noted in Footnote 54, only one additional restriction is required, since it should be the case that the model with monitoring has an identical calibration as the model without monitoring as $\kappa^* \to 0$. However, because a continuum of combinations of μ_κ and σ_κ are consistent with the average flow value of non-employment used to calibrate the model without monitoring, I describe this second step as calibrating both parameters using two restrictions, one of which is the restriction already imposed.

calculated in the main text, based on CPS data from the early 2000s in Stevenson (2009).

D Frictionless model

This appendix characterizes the frictionless limit of the model in Section 3 that I use for the quantitative analysis in Section 5.[57] When the aftermarket is frictionless, provided $\theta < 1$, all firms match immediately, and thus $\hat{q}_t = q_t^n = q_t^o = q_t = 1$.

D.1 Unemployment

Because firms are guaranteed to match by the end of the period, $m_t = v_t$, which implies that the law of motion for unemployment in (19) can be written as

$$u_{t+1} = u_t + \delta(1 - u_t) - (1 - \delta)v_t. \tag{D.1}$$

D.2 Vacancies

Next, combining the vacancy value functions in (38), (39), and (40) with the entry condition in (37) and the fact that $q_t = 1$ implies $v_t^n = v_t$ via (20),

$$v_t = \eta_t G\left(-c + \beta(1-\delta)\left[y_{t+1} - w_{t+1} - \tau_{t+1} + c + G^{-1}\left(\frac{v_{t+1}}{\eta_{t+1}}\right)\right]\right). \tag{D.2}$$

D.3 Monitoring

From (34), the optimality condition for monitoring is

$$\kappa_t^* = \beta(1-\delta)(p_t^m - p_t^w)\mathbb{E}\left[W_{t+1} - U_{it+1}\right]. \tag{D.3}$$

Using the Nash bargaining condition in (43) and the vacancy value functions in (38), (39), and (40),

$$\mathbb{E}[W_{t+1} - U_{it+1}] = \tfrac{\chi}{1-\chi}(y_{t+1} - w_{t+1} - \tau_{t+1} + c) \tag{D.4}$$

which can be used to eliminate $\mathbb{E}\left[W_{t+1} - U_{it+1}\right]$ from the optimality condition in (D.3), yielding

$$\kappa_t^* = \beta(1-\delta)(p_t^m - p_t^w)\tfrac{\chi}{1-\chi}\left(y_{t+1} - w_{t+1} - \tau_{t+1} + c\right). \tag{D.5}$$

Finally, when $F(\kappa_t^*)u_t < v_t$,

$$p_t^m - p_t^w = \frac{1 - v_t/u_t}{1 - F(\kappa_t^*)} \tag{D.6}$$

and when $F(\kappa_t^*)u_t > v_t$,

$$p_t^m - p_t^w = \frac{v_t/u_t}{F(\kappa_t^*)}. \tag{D.7}$$

[57] Note that I index η by t in order to indicate that I will treat this as an exogenous variable in the quantitative analysis of Section 5.

Combining the preceding, the optimal monitoring condition can be written as

$$\kappa_t^* = \begin{cases} \beta(1-\delta)\left(\frac{1-v_t/u_t}{1-F(\kappa_t^*)}\right)\frac{\chi}{1-\chi}\left(y_{t+1} - w_{t+1} - \tau_{t+1} + c\right) & \text{if} \quad F(\kappa^*) \leq v_t/u_t \\ \beta(1-\delta)\left(\frac{v_t/u_t}{F(\kappa_t^*)}\right)\frac{\chi}{1-\chi}\left(y_{t+1} - w_{t+1} - \tau_{t+1} + c\right) & \text{if} \quad F(\kappa^*) > v_t/u_t. \end{cases} \tag{D.8}$$

D.4 Wages and taxes

Equations (D.1), (D.2), and (D.8) are similar to (50)-(52) in the main text, except that they depend on the lump-sum tax τ_t (required to pay for benefits) and the wage w_t, both of which are endogenous variables. The final step is to show that these two variables can be written as functions of u_t, v_t, and κ_t^*.

In the case of wages, the worker's value functions must be used. Specifically, note that it is possible to use (30)-(33) together with the fact that when $q_t = 1$, $p_t = \theta_t = v_t/u_t$, to write

$$\mathbb{E}[W_t - U_{it}] = w_t(1-b) - \textstyle\int_{\kappa_t^*}^{\infty} \kappa dF(\kappa) + \beta(1-\delta)(1 - v_t/u_t)\mathbb{E}[W_{t+1} - U_{it+1}]. \tag{D.9}$$

Eliminating $\mathbb{E}[W_{t+1} - U_{it+1}]$ from the right-hand side using (D.3), eliminating $\mathbb{E}[W_t - U_{it}]$ from the left-hand side using (D.4) (evaluated at time t), and defining $\tilde{\tau}_t \equiv \tau_t/w_t = \frac{bu_t}{1-u_t}$ (which makes use of (44)),

$$\frac{\chi}{1-\chi}\left(y_t - w_t\left(1+\tilde{\tau}_t\right) + c\right) = w_t(1-b) - \textstyle\int_{\kappa_t^*}^{\infty} \kappa dF(\kappa) + \kappa_t^*\left(\frac{1-v_t/u_t}{p_t^m - p_t^w}\right) \tag{D.10}$$

which, using the expression for $p_t^m - p_t^w$ in (D.6) and (D.7), can be solved for w_t as a function of u_t, v_t, and κ_t^* as desired:

$$w_t = \omega(\kappa_t^*, u_t, v_t) = \begin{cases} \dfrac{\frac{\chi}{1-\chi}(y_t + c) + \int_{\kappa_t^*}^{\infty}\kappa dF(\kappa) - \kappa_t^*(1 - F(\kappa_t^*))}{1 - b + \frac{\chi}{1-\chi}\left(\frac{1-(1-b)u_t}{1-u_t}\right)} & \text{if} \quad F(\kappa^*) \leq v_t/u_t \\ \dfrac{\frac{\chi}{1-\chi}(y_t + c) + \int_{\kappa_t^*}^{\infty}\kappa dF(\kappa) - \left(\frac{u_t - v_t}{v_t}\right)\kappa_t^* F(\kappa_t^*)}{1 - b + \frac{\chi}{1-\chi}\left(\frac{1-(1-b)u_t}{1-u_t}\right)} & \text{if} \quad F(\kappa^*) > v_t/u_t. \end{cases} \tag{D.11}$$

In the case of taxes, it is possible to use the preceding together with (44) to write

$$\tau_t = \frac{\omega(\kappa_t^*, u_t, v_t)bu_t}{1-u_t}. \tag{D.12}$$

Thus, given (D.11) and (D.12), equations (D.1), (D.2), and (D.8) fully characterize dynamics in the frictionless model.