Labor Demand Management in Search Equilibrium

Cynthia L. Doniger
Federal Reserve Board

David López-Salido
Federal Reserve Board

June 9, 2016

Abstract

A substantial body of economic literature shows that high unemployment, falling wages, and reduced economic activity can have lasting consequences: hysteresis. We model hysteresis as resulting from a coordination failure among atomistic firms and workers in a frictional labor market that features random search, *ex-ante* wage commitments, and the possibility of worker non-participation. This coordination failure results in a continuum of possible (fragile) equilibria with high-labor-demand equilibria welfare dominating low-labor-demand equilibria. We then introduce changes in labor productivity and show that if transition between equilibria is governed by dynamic best-response then endogenous wage rigidity and persistent changes in labor force participation will arise in response to transitory movements in labor productivity. The model is consistent with a host of medium-run and business cycle facts following a severe recession: outward shifts in the Beveridge curve, both jobless and wageless recoveries, and a reduction in the sensitivity of wages to unemployment. Furthermore, expansions are insufficiently robust and in this sense recessions are scarring. This scarring effect can be mitigated by a stimulus policy in the form of a state-dependent minimum wage policy which facilitates coordination on higher output equilibria.

JEL Classification: E24, J42
Keywords: Hysteresis, Coordination Failure, Wage Rigidity, Labor Force Participation, "Job-less" recovery, "Wage-less" recovery.

Preliminary. We thank Susanto Basu, Marcus Hagedorn, and Benjamin Johannsen for comments and Jason Sockin and Carter Bryson for excellent research assistance. We also thank participants in seminars at Cambridge University, Bank of Italy, EIEF Rome, ECB, Boston College, Philly Fed-Penn, and BIS. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System. Comments and suggestions can be directed to `cynthia.l.doniger@frb.gov` and `david.lopez-salido@frb.gov`.

1 Introduction

A substantial body of economic literature shows that high unemployment, falling wages, and reduced economic activity can have lasting consequences. As Blanchard and Summers (1986a, 1987) suggested, hysteresis sets in. We examine these set of issues by building a model in which there is a potential coordination failure in wage setting and a corresponding multiplicity of equilbria. Our economy features "fragility" in the sense that movements in labor productivity (stemming from TFP shocks, aggregate demand shocks, or other sources) can induce persistent movements between the many equilibria generated by an upward sloping labor demand curve.

We begin by considering a stylized two-player game. One worker and one firm may match and produce. Prior to matching, however, each party has a variety of considerations. Before matching the worker will draw a flow value of leisure from a known distribution. Conditional on the realization of the flow value of leisure, but before matching, the worker may decide to exit the game. We denote this decision as non-participation in the labor force. Also before matching the firm must commit to a posted wage without knowledge of the worker's realization of the flow value of leisure. We show that this game exhibits positive spillovers and strategic complementarities as in Cooper and John (1988).

Positive spillovers and strategic complementarities stem from a thick market externality denied to the firm through the worker's participation threshold and a pecuniary externality denied to the worker through the ex-ante wage posting. An increase in the posted wage–which is here the strategy of the firm–increases the payoff to the worker for all inframarginal draws of the value of leisure through the pecuniary externality and increases the workers optimal participation threshold–the strategy of the worker. Conversely, an increase in the participation threshold–the strategy of worker–may increase the payoff to the firm via a thick-market externality and increases the optimal posted wage–the strategy of the firm. We specify an interval of wage and participation pairs on which a continuum of self-confirming rational expectations equilibria exist. Following the results of Cooper and John (1988), these can be welfare ranked with high-wage and high-participation equilibria dominating low-wage, low-participation equilibria.

The stylized assumptions of this two player game can be micro-founded by embedding the game in a frictional labor market in which a continuum of massless workers meet with a continuum of massless firms through a process of pair-wise random search. This supposition of a frictional labor market endows the firm with the studied market power in wage setting (monopsony) even for arbitrarily small market frictions (Diamond, 1971). Assuming massless agents and random search renders each worker (firm) unable to affect the average participation threshold (wage) through unilateral deviation. This supports modeling these aggregates as exogenous to each decision maker's problem. We close the model with the typical free-entry condition into vacancy posting. This provides a map from wages and participation levels to market tightness and unemployment rates.

Thus, in the full model, we are able to make assertions about the implications of moving between equilibria for tightness and unemployment in addition to the implications for output, participation, and wages already present in the simple two-player game. We follow a large part of the search and matching literature and assume the matching function exhibits constant returns to scale. It is important to note that for us this assumption is orthogonal to the assumptions necessary for multiplicity, unlike Diamond (1982). Still, at cyclical frequencies match *efficiency* may change in our model. This is in line with the observation that the Beveridge curve shifts north-east following particularly severe recessions.

Finally, we subject this frictional labor market to stochastic shocks to labor productivity and impose the restriction that any movements between equilibria are governed by *best-response dynamics* as in Vives (1990, 2005); Cooper (1994). We show that a corollary of the multiplicity of equilibria is that there are many productivity levels for which the same wage and participation threshold pair is an element of the equilibria. As a result, shocks within this range do not induce deviation from the status quo wage and participation threshold: an inaction range in which wages and participation are rigid. On a cyclical frequency two important implications are: (1) contractions in the participation rate are associated with decreases in matching efficiency. This may lead to persistently elevated levels of unemployment. (2) Wages exhibit endogenous stickiness. This may lead to a flatter relationship between the unemployment rate and wages when exiting a contraction than when entering a contraction. This enables us to generate both a "jobless-recovery" and a "wageless-recovery" following severe economic contractions.

Further, recessions in our model are costly in terms of their "scarring" effects: following a severe contraction, economic activity in the post-recession economy is for a time less robust than in the pre-recession economy under identical fundamentals. That is, for an interval of marginal product of labor above the recession level trough, labor force participation, wages, and output are all lower than under identical values of marginal product of labor pre-recession. This is because, in our model, the recent history of shocks is also a determinant of equilibrium selection. When the distribution from which workers' flow values of leisure are drawn has no upper bound, expansions are inefficiently robust regardless of their length or steepness. However, the policymaker can intervene to facilitate coordination to welfare dominating equilibria. The policymaker's lever is an appropriately chosen minimum wage. This can be used to implement the constrained efficient equilibrium at any time.

Our theory accords well with three stylized facts in the medium-run data. (1) Over the last fifty years, the U.S. economy has witnessed a persistent downward trend in participation in the labor force among prime-age males (e.g., Juhn et al. (1991), Murphy and Topel (1997), Juhn et al. (2002), and Elsby and Shapiro (2012)). The last recession and the subsequent slow recovery have formed the most recent piece of evidence of this phenomenon. At the height of the Great Recession, nonparticipation among white males with between 1 and 30 years of labor force experience reached nearly 12 percent, compared with below 6 percent in the early 1970s. (2) Over the same horizon, the aggregate matching function has become less efficient (e.g. Barnichon and Figura (2015); Diamond and Şahin (2015)) with a particularly precipitous drop following the financial crisis. Our theory offers declining labor force attachment as a complementary theory to labor market mismatch in accounting for this decline. (3) Again over the same horizon, labor's share of output has eroded (Elsby et al., 2013; Karabarbounis and Neiman, 2014).

Our paper also fits in an old and newly active literature regarding "fragile" labor market equilibria (Summers, 1992, pp. 332). These papers respond to the observation that seemingly small and/or seemingly transient macroeconomic shifts occasionally yield large and persistent shifts in the labor market. Stagnation of European labor markets in the 1980's and the "jobless recoveries" experienced by the United States after the 2001 and 2008 recessions are prime examples. These theories can be broken into two broad categories: (1) those that revolve around a bifurcation point: e.g. Diamond and Fudenberg (1989); Albrecht et al. (2013); Golosov and Menzio (2015) and Beaudry et al. (2015); and (2) those that posit criteria for switching between different equilibria: e.g. Blanchard and Summers (1986b); Eeckhout and Lindenlaub (2015); Kaplan and Menzio (2013) and Schaal and Taschereau-Dumouchel (2015).[1]

In the first type of paper, periodic crises are a permanent feature of equilibria characterized as orbits rather than a single steady state point. As such, these periodic crises sustain higher welfare equilibria in the non-crisis periods than would otherwise be attainable. For example, the "purge" period in Kaplan and Menzio (2013) sustains high effort in all other periods. A drawback to such theories, which Beaudry et al. (2015) are working to remediate by integrating stochastic aggregate shocks, is the rigidly defined periodicity of the crisis.

In the second type of paper, multiplicity of equilibria yield regions of downward-sloping labor supply, as in the insider-outsider model of Blanchard and Summers (1986b), or upward-sloping labor demand, as in the thick-market externality of Diamond (1982). Evidence that wages have stagnated in the episodes in question suggests that a theory of upward-sloping labor demand is more in line with the evidence. Our paper takes the intuition of Diamond's model and imbeds it in a richer game in which we can analyze the co-movement of labor

[1] These are certainly non-exhaustive lists.

productivity, unemployment, and total output, as well as labor force participation and wages.

Our setup can also be used to generate "animal spirits" type business cycles by means of sun-spots that induce players to correlate expectations on periodically more or less efficient equilibria. The distinction between history dependence and the role of expectations in selecting equilibria has been studied in the macroeconomics literature, with a focus on rational expectations. See, for example, Krugman (1991). The indeterminacy highlighted could be resolved by refining equilibria via global games (Morris and Shin, 2000). However, such refinement yields a one-for-one mapping between fundamentals and equilibrium. Refining instead only the path of equilibrium using the best-response dynamics approach of Vives (1990, 2005); Cooper (1994) provides for history dependence in equilibrium selection and allows our model to generate hysteresis.

Although this is not a focus of our paper, endogenous wage stickiness due to our coordination failure also speaks to the Shimer (2005) puzzle. As such, it relates to a larger literature that seeks to induce amplification through persistence in the wage level: for example Hall (2005), Gertler and Trigari (2009), and Kennan (2010). We differ from this literature in our focus on producing hysteresis as opposed to only amplification. In particular, a strict focus on amplification will yield "wage-less" but "job-full" recoveries following particularly severe contractions. Our model can generate both a "job-less" and "wage-less" recovery–that is a recovery in which unemployment remains persistently high even after labor productivity has recovered and a recovery in which unemployment falls to unprecedented lows before wages recover, respectively. An implication is that the social planner can improve total welfare by *raising* the wage.

The rest of the paper is structured as follows. The next section presents empirical evidence that motivates key features of the model. Section 3 describes the benchmark model and establishes that each player's strategy imposes positive spillovers on and is a strategic complement to the others. Section 3.2 houses the basic intuition of the baseline model in a frictional labor market and argues that search friction microfounds the assumptions of the baseline model. Section 5 expands the model to admit stochastic shocks to labor productivity and establishes a range of such shocks for which no player changes strategy. Section 6 investigates how this region of inaction interacts with transitory labor productivity shocks to produce business cycles typified by swift contraction and sluggish recovery. Section 7 discusses some policy implications. Section 8 tests model implications in United States labor market data. Section 9 concludes and offers suggestions of further avenues of investigation.

2 Motivating Empirical Evidence

2.1 Medium Run Variation in Labor Force Participation

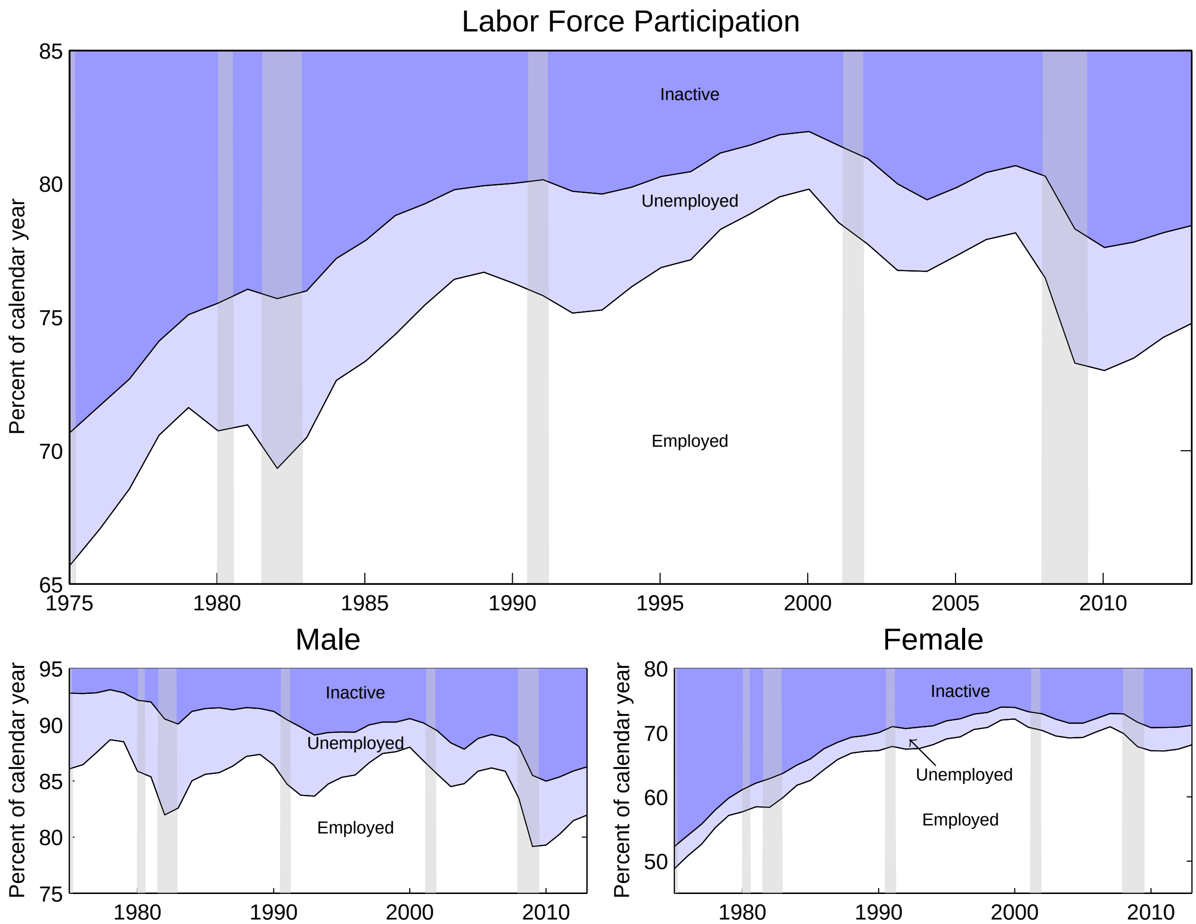
Using data from the Annual Demographic File of the Current Population Survey, Figure 1 plots the average fraction of weeks that healthy, non-disabled civilians with between 1 and 30 years of experience who were not engaged in education spent in employment, unemployment and nonparticipation in each year since 1975.[2] The figure suggests a key feature: an apparently cyclical trend in non-employment masks an (alarming) medium-run trend toward non-participation in the labor force.[3]

The measured fraction of yearly weeks unemployed for men has actually declined since the mid-1970's and exhibits sharp cyclical movements; however, non-participation has risen with a dramatic secular trend from just over 6 percent in the mid-1970's to over 11 percent in 2014. In other words, since the early 70s, males in US economy have exhibited a medium-run trend toward lower levels of labor force participation. For females a similar trend begins in the 2000s. This trend exhibits an asymmetric cyclical responses to positive (expansions) and negative (recessions) shocks. Nonparticipation in the labor force accelerates during recessions

[2]Similar stylized facts about labor force participation were initially documented by Juhn et al. (1991), and subsequently by Juhn et al. (2002) and Elsby and Shapiro (2012). We extend the data to include the recent recession and isolate our results from the coincident and comparatively acyclical trend in disability.

[3]This is most pronounced for males. A secular trend toward participation among females obscures this feature in that population in the last century; however, since 2000 the pattern for females mirrors that of males.

Figure 1: Employment, Unemployment, and Inactivity.
(Sample Period: 1975–2013)



Source: Authors' calculations based on the Annual Demographic File of the Current Population Survey following the methodology of Juhn et al. (2002) and Juhn et al. (1991).
Sample: Civilians with 1 to 30 years of potential experience. Individuals who report being students, retired, or ill/disabled are excluded.

(sharp contraction) and does not reverse during expansions to the pre-recessions levels (slow recovery): while unemployment falls during expansions, participation rates hardly recover.

Table 1, taken from Juhn et al. (2002) and extended to reflect recent data, records unemployment and non-participation rates for males in business cycle peaks and troughs since 1975. The pattern is clear: labor force participation falls during contractions and does not recover sufficiently during expansions.

Table 1: Unemployment, Nonparticpation, and Nonemployment during Business-Cycle Peaks and Troughs, 1975-2013.

Period	Phase of business cycle[b]	Labor force status (percent of calendar year[a]) unemployed	inactive	non-employed[c]	Δ unempoyed (%)	Δ inactive (%)
1967-69	Peak	2.2	4.1	6.3		
1971-72	Trough	4.5	4.9	9.4	2.3	0.8
1972-73	Peak	3.8	5	8.8	-0.7	0.1
1975-76	Trough	6.9	5.6	12.4	3	0.6
1978-79	Peak	4.3	5.9	10.2	-2.5	0.3
1982-83	Trough	9.0	6.3	15.2	4.6	0.4
1988-89	Peak	4.3	6.7	11	-4.7	0.5
1991-92	Trough	6.3	7.5	13.8	2	0.8
1999-00	Peak	3.0	8	11	-3.3	0.5
2002-03	*Trough*	*3.5*	*10.2*	*13.7*	*1.0*	*0.9*
2005-06	*Peak*	*2.8*	*10.9*	*13.7*	-0.7	*0.7*
2008-09	*Trough*	*4.4*	*12.0*	*16.4*	*1.6*	*1.1*

Source: Juhn et al. (2002). Authors' additional calculations based on the Annual Demographic Files of the Current Population Survey appear in italic.

[a] Number of weeks in indicated status divided by 52.

[b] Peaks and troughs are defined as years with minimal and maximal unemployment within each cycle.

[c] Unemployed plus nonparticipation details may not sum to totals because of rounding.

2.2 Shifting Beveridge Curve and Declining Match Efficiency

The previous analysis is often discussed in the context of shifts of the Beveridge curve – the relationship between unemployment rates versus vacancy rates. In particular, outward shifts in the Beveridge curve have been common occurrences during U.S. recoveries, and among the potential sources of shifts are changes in labor force participation. Although these elements are connected, the link between the two remains loose. Following the work by Diamond and Şahin (2015) and looking at 60 years of data, there is a notable outward shift in the Beveridge curve after the maximum unemployment rate is reached (Table 2): In seven of the eight completed business cycles, the Beveridge curve shifted outward; and in three of these cycles, the unemployment rate undershoots below its pre-recession level in the next expansion, while in four it did not. These observations suggest an important asymmetry: by itself, an outward shift in the Beveridge curve does not predict how low unemployment

Table 2: Beveridge Curve shifts, minimum unemployment rate, and length of expansions for different business cycles.

Max-U	Experiences Shift	Min U (Recovery)<Min U(Previous Expansion)	Length of Expansion
Q3-1954	YES	NO	11
Q2-1958	YES	NO	5
Q2-1961	YES	YES	32
Q3-1971	YES	NO	10
Q2-1975	YES	NO	17
Q4-1982	YES	YES	26
Q3-1992	YES	YES	34
Q2-2003	NO	NO	15
Q4-2009	YES	N/A	N/A

Source: Diamond and Şahin (2015).

gets during the recovery. Underlying this asymmetry rests the duration of the expansion. As Diamond and Şahin (2015) noted, longer expansions are the ones in which the unemployment rate goes below its pre-recession trough.

The left panel of Figure 2 graphically supports the previous description by showing the relationship between vacancies and unemployment changes in two different U.S. cyclical episodes. In fact, through this lens the recent recession resembles other deep recessions and slow recoveries, such as the experience during and after 1973.

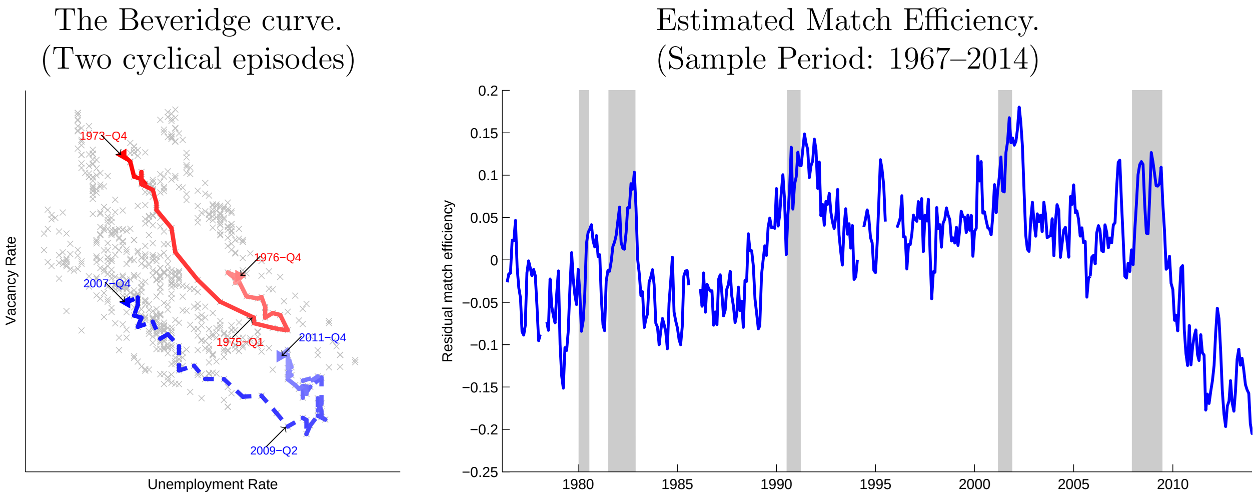
An alternative way to capture shifts in match efficiency is to examine the residual in an estimation of the matching function, as in Barnichon and Figura (2015). The right panel of Figure 2 plots the residual from an estimation of the matching function:

$$\ln(Job\ Finding\ Rate) = \ln(Match\ Efficiency) + (1-\eta)\ln(Labor\ Market\ Tightness) + \varepsilon$$

where η is the elasticity of the job finding rate with respect to unemployment.[4] The residual clearly exhibits predictability at the cyclical frequency, especially following the Great Recession.

[4]The specification of the matching function will be discussed in greater detail in sections 4 and 8.1.

Figure 2: Changes in Match Efficiency.



Source (left): Bureau of Labor Statistics, Current Population Survey; Conference Board, Help Wanted Index and Help Wanted Online Index; Bureau of Labor Statistics, Job Openings and Labor Turnover Survey. Vacancies are constructed from the Conference Board data as in Barnichon (2010) prior to 2001 and follow the Job Opening and Labor Turnover Survey thereafter. The figure is constructed as in Diamond and Şahin (2015)
Source (right): Authors calculations based on Bureau of Labor Statistics, Current Population Survey; Conference Board, Help Wanted Index and Help Wanted Online Index; Bureau of Labor Statistics, Job Openings and Labor Turnover Survey. Vacancies are constructed from the Conference Board data as in Barnichon (2010) prior to 2001 and follow the Job Opening and Labor Turnover Survey thereafter. The figure is constructed as in Barnichon and Figura (2015).

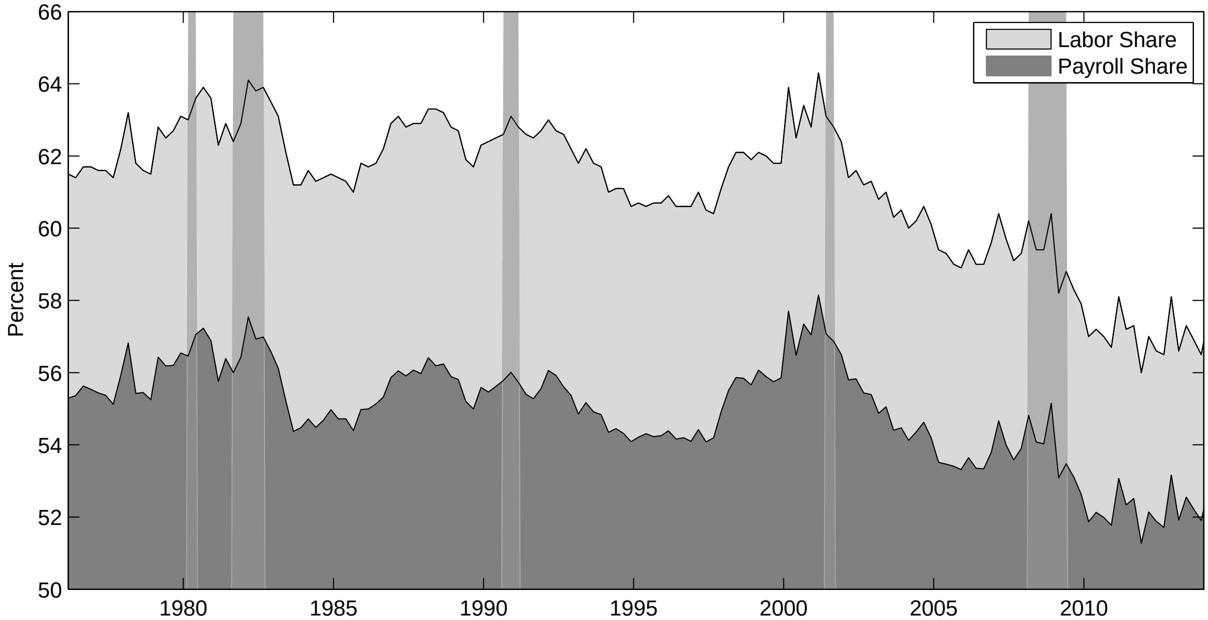
2.3 Medium Run Variation in Labor Share

A third piece of evidence we appeal to in order to argue in favor of a labor demand side explanation for the observed trends is the fall in labor share over the same horizon. Figure 3 plots the Bureau of Labor Statistics official labor share (fraction of value added paid as compensation of labor) as well as "payroll share." Payroll share is defined as in Elsby et al. (2013) as the ratio of wages and salaries to value added. Payroll share thus is not biased by potential changes in the ratio of proprietors income to employees income that may have occurred over this time horizon.[5]

Both labor and payroll share exhibit a clear pattern: the share increases interring a contraction and in the subsequent expansion the share falls below the pre-contraction level. We take the first of these as evidence of a compensation bias: low output workers, who are

[5] Elsby et al. (2013) demonstrates that labor share is inflated through the 1980's due to the Bureau of Labor Statistics algorithm for imputing proprietors income.

Figure 3: Labor Share and Payroll Share.
(Sample Period: 1976–2015)



Source: Bureau of Labor Statistics, Bureau of Economic Analysis, and author's calculations following Elsby et al. (2013).

likely also low wage, are more likely to be separated during a contraction. We take the second of these to be evidence of a structural trend related to the fall in labor force participation and in match efficiency.

The goal of this paper is to present a theoretical framework explains these findings. In so doing we provide a firm-side (demand for labor) mechanism that underlies movements in labor force participation and shifts in the Beveridge curve. Our model posits that transitory contractions in labor productivity lead to scarring effects on the labor demand side. This leads to weak labor demand even after labor productivity has recovered and provides an explanation for these medium-run trends.

3 A Stylized Two Player Game

Note that solving the two player game requires positing a fixed expectation for the firm's posted wage on the worker's behalf–we will call the worker's expectation of the wage w_0–and for the worker's participation threshold on the firm's behalf–we will call the firm's expectation of the threshold r_0. In this section we remain agnostic as to how such expectations are formed and focus only on demonstrate that there exist multiple self-confirming rational expectations equilibria. In following sections we posit an expectation formation process that depends on the realized history of aggregate shocks.

We begin by considering a stylized one-worker-one-firm labor market. The worker's objective is to maximize her compensation. The firm's objective is to maximize profit.

3.1 Worker

The worker will draw a value of leisure, b, from a known distribution, $H(b)$, with density $h(b)$ and with support $[\underline{b}, \bar{b}]$ with $\bar{b}$ potentially infinite, and may subsequently be presented with a wage offer. Prior to this draw the worker must choose a strategy for participation in the labor market. As is typical in such a problem, the worker's strategy takes the form of two thresholds: a reservation wage and a participation threshold. Clearly, any wage offer below the flow value of leisure will be rejected. Thus, the reservation wage is equal to the realized flow value of leisure. The more interesting problem is to establish the participation threshold. Given the reservation wage and participation threshold strategy the worker may end up in one of four states: employed, involuntarily unemployed, voluntarily unemployed, or inactive. Employment occurs if the worker participates, matches, and accepts the wage offer. Involuntary unemployment occurs if the worker participates, the posted wage offer would be acceptable, but no match occurs. Voluntary unemployment occurs if the worker participates but the posted wage offer is not acceptable. Inactivity occurs if the worker does not participate.

We allow for the following participation threshold strategy: below some threshold r participate, else play a mixture on participate and non-participate with the weight on non-participant being $i \in [0,1]$. Note, when the reservation threshold is set as the best response to the expected posted wage offer of the firm, w_0, the worker is indifferent between voluntary unemployment an inactivity. As a result, we view i as a mixed strategy with these pure strategies in the support. In later sections we will consider policies that break the indifference between these states.[6] Having chose her strategy the worker will meet with the firm with exogenous probability $(1-u)$ if participating. After meeting with the firm the worker accepts or rejects the posted wage offer, w_0.[7]

Let $V^W(r, w_0)$ be the expected payoff to a representative worker of choosing threshold r when the expected wage choice of the firm is w_0. Let $r^*(w_0)$ be the best response of the representative worker when the expected wage choice of the firm is w_0. We can write the payoff of selecting threshold r given the expected posted wage, w_0. It is easiest to do this

[6] While a pure strategy of non-participate may seem more reasonable on this region we note that for realizations of $b \geq r$ the worker is indifferent between non-participation and participation coupled with rejection of all offers (perpetual unemployment). Further, we will see that the stark assumption that workers with high values of nonemployment play a pure strategy on non-participate produces undesirable implications in the stochastic economy: eventual convergence to a zero-output steady state. In later sections, we also explore the possibility that the social planner may have a policy instrument–search contingent unemployment insurance–that breaks the indifference between "participate" and "non-participate" in favor of "participate". We will see that such a policy will exacerbate unemployment during contractions but will also lead to more robust expansions.

[7] Note, the unemployment probability will be pinned down in equilibrium via the aggregate matching function. Importantly, however, the unemployment rate is independent of each worker's participation choice (since each is massless) and thus, as we are about to see, each worker's choice of participation threshold is independent of the unemployment rate.

separately for the case $r \le w_0$ and for the case $r > w_0$. We begin with $r \le w_0$:

$$\begin{aligned} V^W(r, w_0) = & \underbrace{(1-u)w_0\big[H(r) + (1-i)[H(w_0) - H(r)]\big]}_{employed} \\ & + \underbrace{u\big[\int_{\underline{b}}^{r} bh(b)db + (1-i)\int_{r}^{w_0} bh(b)db\big]}_{involuntarily\ unemployed} \\ & + \underbrace{(1-i)\int_{w_0}^{\bar{b}} bh(b)db}_{voluntarily\ unemployed} \\ & + \underbrace{i\int_{r}^{\bar{b}} bh(b)db.}_{inactive} \end{aligned} \tag{1}$$

The value is the sum of the payoff in the case of employment, involuntary unemployment, voluntary unemployment, and inactivity, weighted according to their likelihood. Employment occurs in two cases. First, the worker may draw $b < r$ triggering participation with certainty–this occurs with probability $H(r)$–and subsequently meet an employer and accept the wage offer-this occurs with probability $(1-u)$. Second, the worker may draw $b \in [r, w_0]$ but still select participation–this occurs with probability $[H(w_0) - H(r)](1-i)$–and subsequently meet an employer and accept the wage offer-this occurs with probability $(1-u)$. In both of these cases the payoff is w_0. Thus we have the first line of equation (1).

Involuntary unemployment also occurs in two cases. First, the worker may draw $b < r$ triggering participation with certainty–this occurs with probability $H(r)$–but subsequently fail to meet an employer -this occurs with probability u. Second, the worker may draw $b \in [r, w_0]$ but still select participation–this occurs with probability $[H(w_0) - H(r)](1-i)$–but subsequently fail to meet an employer-this occurs with probability $(1-u)$. In these cases the payoff is the random draw of b and the expected payoff is the conditional expectation of b: that is $\int_{\underline{b}}^{r} bh(b)db$ and $\int_{r}^{w_0} bh(b)db$, respectively. Thus we have the second line of equation (1).

Voluntary unemployment occurs if the worker draws $b > w_0$ and she randomizes to participation–this occurs with probability $(1-i)[1-H(w_0)]$. If this worker subsequently meets with a firm she rejects the wage offer, thus her payoff in this case is invariant to u. The payoff is again the conditional expectation of b: that is $\int_{w_0}^{\bar{b}} bh(b)db$. Thus we have the third line of equation (1).

Finally, inactivity occurs if the worker draws $b > r$ and she randomizes to non-participation. This occurs with probability $i[1-H(r)]$. Again the payoff is the conditional expectation of b: that is $\int_r^{\bar{b}} bh(b)db$. Thus we have the final line of equation (1).

Differentiating, using Leibnitz rule, one can easily show that $\frac{dV^W(r,w_0)}{dr} \geq 0$ whenever $r \leq w_0$. The value function and derivative may appear complicated but the intuition is simple: if the threshold is low enough that there is a positive probability that an acceptable wage offer would be forgone due to non-participation then the threshold is too low. In other words, the participation threshold should be at least as high as the expected wage.

Similarly we can write the value function for the case $r > w_0$:

$$
\begin{aligned}
V^W(r,w_0) = & \underbrace{(1-u)w_0\big[H(w_0)\big]}_{employed} \\
& + \underbrace{u\big[\int_{\underline{b}}^{w_0} bh(b)db\big]}_{involuntarily\ unemployed} \\
& + \underbrace{\int_{w_0}^{r} bh(b)db + (1-i)\int_r^{\bar{b}} bh(b)db}_{voluntarily\ unemployed} \\
& + \underbrace{i\int_r^{\bar{b}} bh(b)db.}_{inactive}
\end{aligned} \tag{2}
$$

Again, the value can be dissected into cases: employment, involuntary unemployment, voluntary unemployment, and inactivity. Employment occurs if the worker draws $b < r$ triggering participation with certainty and subsequently meets an employer and accept the wage offer–this occurs with probability $(1-u)H(w_0)$. The binding consideration here is wage offer acceptance as opposed to the $r < w_0$ case in which the binding consideration was participation. The payoff is w_0. Thus we have the first line of equation (2).

Involuntary unemployment if the worker draws $b < r$ triggering participation with certainty, subsequently fails to meet an employer, and would have accepted an employment opportunity if offered-this occurs with probability $uH(w_0)$. Again, the binding consideration is wage offer acceptance as opposed to the $r < w_0$ case in which the binding consideration was participation. The payoff is the random draw of b and the expected payoff is the conditional expectation of b: that is $\int_{\underline{b}}^{w_0} bh(b)db$. Thus we have the second line of equation (2).

Voluntary unemployment occurs if the worker draws $b > w_0$ and selects participation–either because $b \in [w_0, r]$ or because $b > r$ and she randomizes to participation-this occurs with probability $[H(r) - H(w_0)] + (1 - i)[1 - H(w_0)]$. If this worker subsequently meets with a firm she rejects the wage offer, thus her payoff in this case is invariant to u. The payoff is again the conditional expectation of b: that is $\int_{w_0}^{r} bh(b)db$ and $\int_{r}^{\bar{b}} bh(b)db$, respectively. Thus we have the third line of equation (2).

Finally, inactivity occurs if the worker draws $b > r$ and she randomizes to non-participation. This occurs with probability $i[1 - H(r)]$. Again the payoff is the conditional expectation of b: that is $\int_{r}^{\bar{b}} bh(b)db$. Thus we have the final line of equation (2).

Differentiating, again using Leibnitz rule, one can easily show that $\frac{dV^W(r,w_0)}{dr} = 0$ whenever $r > w_0$. Again, the calculus is tedious but the intuition simple. All participation thresholds strictly above the expected wage level yield the same payoff since the worker is indifferent between voluntary unemployment and inactivity when her draw of the flow value of leisure exceeds the wage level.

We have $\frac{dV^W(r,w_0)}{dr} \geq 0$ whenever $r \leq w_0$ and $\frac{dV^W(r,w_0)}{dr} = 0$ otherwise. Thus the optimal participation threshold strategy is the correspondence bounded below by:

$$r^*(w_0) = \begin{cases} \underline{b} & \text{if } w_0 < \underline{b} \\ w_0 & \text{if } w_0 \in [\underline{b}, \bar{b}] \\ \bar{b} & \text{if } w_0 > \bar{b} \end{cases} \tag{3}$$

Proposition 1. *This game exhibits* ***positive spillovers*** *and (weak)* ***strategic complementarities*** *for the worker.*

Proof. We follow Cooper and John (1988). We define ***positive spillover*** for player j as the case in which an increase in the other player's strategy increases the payoff to player j. Now simply note that $\frac{dV^W(r,w)}{dw} = H(r) > 0$. In words, an increase in the posted wage–which we will see is the strategy of the firm–increases the payoff for the worker regardless of worker's participation threshold strategy. Also we define ***strategic complementarities*** for player j as the case when an increase in the other player's strategy increases the best response of player j. Now again simply note that $\frac{dr^*(w)}{dw} = 1 > 0$. In words, an increase in the posted wage–which we will see is the strategy of the firm–increases the lower bound on the optimal participation threshold strategy of the worker.

Firm

The firm posts a wage offer ex-ante, w, and may meet the worker with probability $(1 - v)$.[8] If the firm successfully hires the worker it will earn rent $p - w$. If the firm and the worker do not meet or if the worker rejects the wage offer then the payoff to the firm is zero.

Let $V^F(w, r_0)$ be the payoff to the firm of posting wage w and let $w^*(r_0)$ be the best response of the firm when the firm's expectation of the worker's threshold for participation is r_0. We can write the payoff to the firm of selecting wage w when the expected participation

[8] Note, the vacancy rate will be pinned down in equilibrium via the aggregate matching function and a free entry condition. Importantly, however, the vacancy rate is independent of each firms's wage and vacancy posting choice (since each is massless) and thus, as we are about to see, each player's choice of participation or wage is independent of the vacancy rate.

threshold of the worker is r_0:

$$V^F(w, r_0) = \underbrace{(1-v)}_{\substack{\text{match}\\ \text{probability}}} \underbrace{(p-w)}_{\substack{\text{revenue}\\ \text{from}\\ \text{successful}\\ \text{hiring}}} \underbrace{\left[\frac{\mathbb{I}_{\{w \le r_0\}} H(w) + (1 - \mathbb{I}_{\{w \le r_0\}})\big[H(r_0) + i[H(w) - H(r_0)]\big]}{H(r_0) + i(1 - H(r_0))}\right]}_{\textit{expected labor supply schedule}}, \tag{4}$$

where $\mathbb{I}_{\{w \le r_0\}}$ is an indicator function equal to 1 if the condition holds. The first term, $(1-v)$, and last term, $(p-w)$, are the probability of meeting the worker and the payoff from successfully hiring, respectively. These do not require much discussion.

The bracketed term contains the core of the problem: this encodes the firm's expectation of the labor supply schedule it faces. That is the probability that a wage offer of w will be accepted if the firm does indeed meet the worker. Notice that the participation threshold introduces a kink in the labor supply at the expected threshold. To the left of the threshold the labor market is "thick": the worker will participate with certainty for draws of the value of leisure in this region. To the right of the threshold the labor market is "thin": the worker participates with some probability $i \le 1$ for draws of the value of leisure in this region. Posting a wage less than or equal to the expected labor force participation threshold, $w \le r_0$, results in hiring a worker with probability $\frac{H(w)}{H(r_0))+i(1-H(r_0))}$. The denominator counts all of the cases in which the worker might participate and meet the firm and the numerator counts the fraction of these cases in which the worker accepts the wage offer. Posting a wage greater than the expected labor force participation threshold, $w > r_0$ results in hiring a worker with probability $\frac{H(r_0)+i[H(w)-H(r_0)]}{H(r_0))+i(1-H(r_0))}$: again, the denominator counts all the cases in which the worker might participate and meet the firm and the numerator counts the fraction of these cases in which the worker accepts the wage offer. So, we can see that the elasticity of hiring with respect to the posted wage above the threshold is strictly less than the elasticity of hiring with respect to the posted wage below the threshold.

The firm's best response, which pins down its optimal ex-ante posted wage, satisfies:

$$w^*(r_0) = \arg\max_w \left\{V^F(w, r_0)\right\}.$$

As is usual for a wage-posting firm, the problem takes the form of monopsony wage setting. The firm understands and internalizes the impact of it's wage choice on labor supply. Our problem is slightly more complex due to the expectation element of the game. As noted, the firm's expectation over the worker's participation threshold imposes a kink in the expected labor supply curve. As a result the best response function–the optimal wage posting strategy of the firm–is defined piecewise on intervals of the expectation for the worker's participation threshold. Our problem now is to determine these intervals and the optimal posted wage schedule within each interval.

A first case is that the firm expects that the participation threshold is high enough that it is non-binding. Call the minimum non-binding threshold r^C. In this case $V^F(w, r_0|r_0 \geq r^C)$ simplifies to $(1-v)(p-w)\frac{H(w)}{H(r_0))+i(1-H(r_0))}$ and we have:

$$w^C = \arg\max_w \left\{V^F(w, r_0|r_0 \geq r^C)\right\} = \arg\max_w \left\{[H(w)(p-w)\right\},$$

where the final equality follows from noting that $(1-v)$ and $H(r_0)+i[1-H(r_0)]$ are exogenous from the perspective of the firm. The first order condition gives:

$$h(w^C)(p - w^C) = H(w^C). \tag{5}$$

The left hand side is the benefit of an incremental increase in the wage around w^C and the right hand side is the cost. Since we have posited that r_0 is non-binding in this case the first order condition holds with equality.[9] Recalling the worker's reservation wage and

[9] We can interpret equation (5) in the usual context of a monopsony: $w^C = p - H(w^C)/h(w^C)$ where $H(w^C)/h(w^C)$ is the monopsony markdown. In this simple two-player game the monopsonistic wage setting of the firm imposes inefficiency: some rent-producing matches are not formed as the firm restricts hiring in order to capture rents through the markdown. We will see, however, that in the full model including

participation threshold policies we have that $r^C = w^C$.[10]

A second case is that the firm expects that the participation threshold is low enough that it is binding. In absence of the constraint the firm would optimize at a wage strictly greater than the expected participation threshold. In other words, the benefit of an incremental increase in the wage in the neighborhood of the participation threshold exceeds the cost in the thick side of the market (taking the derivative from the left):

$$h(r_0)\,(p - r_0) > H(r_0).$$

There are two subcases. The marginal cost may also fall short of the marginal benefit of hiring in the neighborhood of the threshold on the thin side of the market. In this case we have the optimization problem:

$$\hat{w} = \arg\max_w \left\{ V^F(w, r_0 | r_0 < r^C) \right\} = \arg\max_w \left\{ [H(r_0) - i[H(w) - H(r_0)](p - w) \right\}$$

In this case the first order condition gives

$$ih(\hat{w})(p - \hat{w}) = H(r_0) + i[H(\hat{w}) - H(r_0)]. \tag{6}$$

Again left hand side is the benefit of an incremental increase in the wage around $\hat{w}$ and the right hand side is the cost.[11] In the first subcase, the marginal benefit falls short of the

labor market frictions the rents that the monopsony expects to capture finance the cost of vacancy posting. Without monopsony rent no vacancies will be posted. Indeed, if one solves a planner's problem for the optimal wage in the frictional labor market without expectation constraints one finds that the monopsony wage choice (wage posting) is the optimal wage (wage contracting arrangement). See for example Acemoglu and Shimer (1999).

[10] For general $H(b)$, w^C is not guaranteed to be unique. A necessary condition for uniqueness is second order condition: $\frac{2h(b)}{p-b} > \frac{d^2H(b)}{db^2}$. This is satisfied in fairly general circumstances – for example the normal distribution when at least half of workers participate. For illustrations we take $H(b)$ to distribute uniform. The appendix characterizes the full set of equilibria when w^C is not unique. The remainder of the body of this paper focuses on the case where the second order condition holds. This is the case under fairly general conditions. In particular, this condition is satisfied by the uniform and Pareto distributions as well as by the normal and logistic distributions when at least half of workers participate in the labor force.

[11] Again, we can interpret equation (5) in the usual context of a monopsony: $\hat{w} = p - [H(r_0) + i[H(\hat{w}) - H(r_0)]]/ih(\hat{w})$ where $H(w^C)/h(w^C)$ is the monopsony markdown.

unconstrained case in proportion to the thinning of the market. The marginal cost, however, falls short by a smaller factor since the incremental increase in the wage must be paid to inframarginal workers on the thick side of the market as well as on the thin side.

In a second subcase, it is also possible that the cost exceeds the benefit of hiring in the entire thin side of the market:

$$ih(r_0)\,(p - r_0) < H(r_0).$$

In this subcase the optimal wage is a corner solution at $w^*(r_0) = r_0$.

What remains is to find the expected participation threshold that is the boundary between these subcases, call it r^L. The boundary will be the expected participation threshold that weakly satisfies the first order condition of the constrained problem:

$$ih(r^L)(p - r^L) = H(r^L) \tag{7}$$

So we have:

$$w^*(w_0) = \begin{cases} \hat{w} & \text{if } r_0 < r^L \\ r_0 & \text{if } r_0 \in [r^L, r^C] \\ w^C & \text{if } r_0 > r^C \end{cases} \tag{8}$$

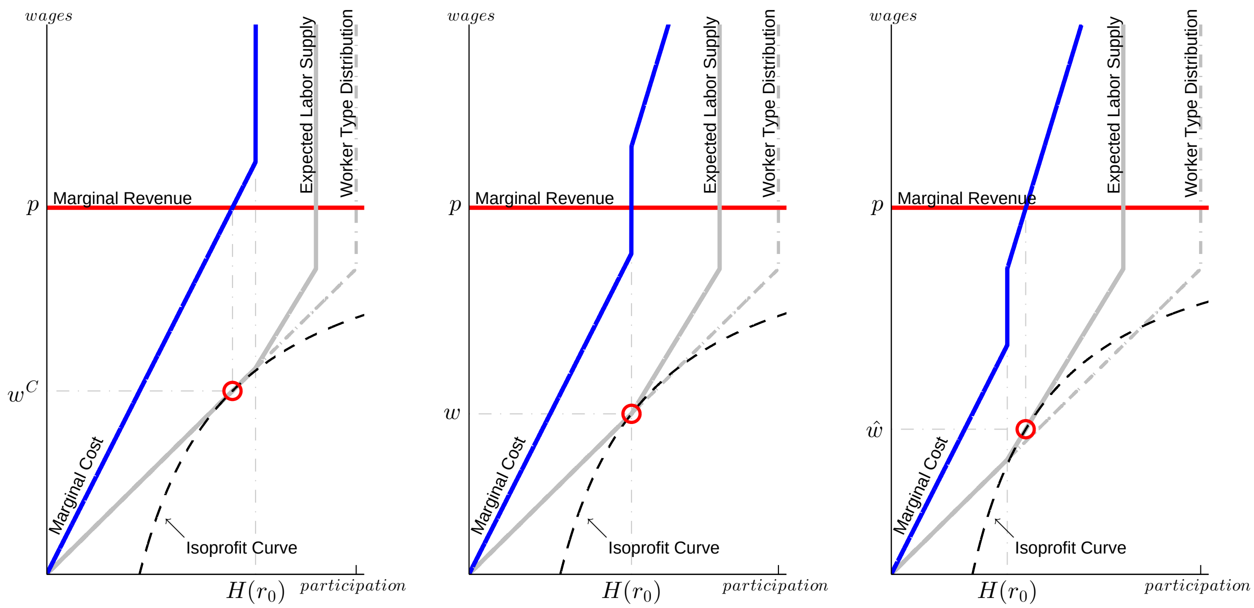
as the best response of the firm.

Again recalling the worker's reservation wage and participation threshold policies we have that $r^C = w^C$ and $r^L = w^L$. For the remaind of the paper we will refer to the interval defined in wage space $[w^L, w^C]$.

Proposition 2. *For aggregate wage levels in the interval $[w^L, w^C]$ the game exhibits **weakly positive spillovers** and **strategic complementarities** for the firm.*

Proof. Again this follows from Cooper and John's (1988) definitions. Note that $\frac{dV^F(w,r)}{dr} =$

Figure 4: Firm's Constrained Wage Choice as the Participation Threshold Declines.



$(p-w)h(r) > 0$ whenever $r_0 \in [w^L, w^C]$. In words, an increase in the participation threshold–which we have seen is the strategy of the worker–increases the payoff for the firm regardless of firms's posted wage strategy in the interval $[w^L, w^C]$. We also have $\frac{dw^*(r)}{dr} = 1 > 0$ whenever $r \in [w^L, w^C]$. In words, an increase in the participation threshold increases the optimal posted wage strategy of the firm. □

Figure 4 gives a graphical representation of the firm's monopsony problem. To simplify the diagram we take $H(w)$ to be uniformly distributed on $[\underline{b}, \bar{b}]$, so $H(w) = \frac{w-\underline{b}}{\bar{b}-\underline{b}}$. Now we can invert to find $w(L|r_0)$, where $L|r_0$ is the labor supply schedule that the firm expects to face. We plot both the underlying distribution of worker types and the labor supply schedule that the firm expects to face given the worker's participation threshold in gray solid and hashed lines respectively.[12] Note that the expected labor supply schedule lies to the north of the underlying distribution of worker types in the region to the east of the quantity of labor supplied at the expected participation threshold, $H(r_0)$. This reflects the thinner market faced by the firm in this region: an incremental increase in the wage results in fewer additional

[12]Note that both schedules are censored and become vertical at $H(\bar{b}) = 1$.

hires to the east of the quantity supplied at the expected participation threshold (that is, the slope of the expected labor supply schedule exceeds that of the underlying distribution of worker types in the region above the quantity of labor supplied at the participation threshold). Now we plot the marginal revenue and marginal cost faced by the firm. Marginal revenue is clear, p. Marginal cost results from solving the firm's optimization for the quantity of labor demanded as a function of wages. Note that to the left of the quantity of labor supplied at the expected participation threshold the marginal cost of an incremental increase in hiring is strictly less than the marginal cost of an incremental increasing in hiring to the right of the quantity of labor supplied at the expected participation threshold. At the quantity of labor supplied at the expected participation threshold the marginal cost is a correspondence: the lower bound being the marginal cost of hiring the $H(r_0) - \varepsilon$ worker and the upper bound being the marginal cost of the $H(r_0) + \varepsilon$ worker. In this region the firm finds a corner solution.

As in the typical monopsony problem we find the optimal quantity of labor demanded at the intersection between marginal revenue and marginal cost. The optimal wage is then found by evaluating the expected labor supply schedule at this quantity of labor. The shortfall between wages and revenue is the monopsony markdown. We also plot the iso-profit curve of the firm. We also plot the maximal iso-profit curve of the firm. This is tangent to the expected labor supply curve at the optimal wage and labor quantity choice. Note that tangency may arise at a corner solution! It is useful to note the geometry of the iso-value curve: a hyperbola, $\frac{\Pi}{H(b)(p-w)}$, where Π is the value at every point along some iso-value curve. The hyperbola has asymptotes at $b = 0$ and $w = p$.

The left panel depicts the unconstrained case: the quantity of labor supplied at the expected participation threshold, $H(r_0)$, exceeds the optimal quantity of labor demanded and the firm's problem reduces to that of a standard monopsony. The firm chooses to post wage w^C and hire in the $H^C) \leq H_0$ fraction of realizations of b where $b \leq w^C$. In the center and right panel the expected participation threshold binds. In the center panel the expected participation threshold falls in the region $[r^L, r^C]$. When this is the case, marginal revenue exceeds marginal cost of hiring an additional worker from the thick side of the market, but it falls short of the marginal cost of hiring an additional worker from the thin side of the marker: profit is maximized at the corner solution, $w^*(r_0) = r_0$. The wage is again found by evaluating the expected labor supply schedule at H_0 and we have $w \in (w^L, w^C)$. In the right panel the participation threshold is low enough that the firm strictly prefers to higher some workers from the thin side of the labor market: marginal revenue and marginal cost intersect in the steep portion of the marginal cost curve. Again, wages are found by evaluating the expected labor supply schedule at the optimal quantity of labor and we have wage $\hat{w}$ and quantity of labor demanded $\hat{H}$.[13] Note, as we established in the proof of proportion 2 $H(r_0)$, as rises between $H(r^L)$ and $H(r^C)$ the posted wage that best responds also rises: illustrating the strategic complementarity.

3.2 Self-Confirming Rational Expectations Equilibria

Definition 1. *A* self-confirming rational expectations equilibrium *of the two-player game is a double –* ***wage level, participation threshold*** *– such that the wage level posted by the firm and participation threshold of the worker are mutual best responses.*

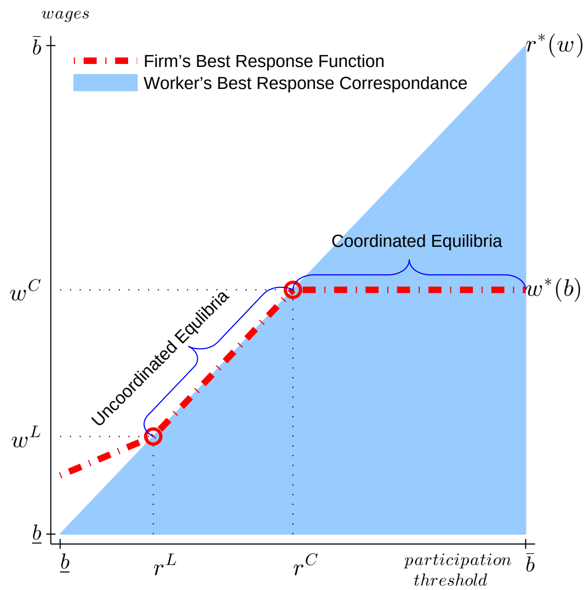
Proposition 3. *A continuum of equilibria exists – wage levels in the interval* $[w^L, w^C]$ *– with higher welfare for higher wage levels.*

Proof. This is a straightforward application of Cooper and John (1988) Proposition 5. □

Proposition 2 enables us to rank equilibria in the two-player game in terms of the output gap defined as the difference between total output at an equilibria with an arbitrary wage level and total output at the constrained efficient equilibrium at wage level w^C.

[13] We will see momentarily that this last case can not be part of a rational expectations equilibrium. The wages proposed at this point exceed the posited participation threshold and the worker would want to instead deviate to a higher threshold in order to accept wages in the interval between the posited threshold and the wage level that constitutes the best response to this threshold.

Figure 5: Mutual Best Responses.



Note: A single coordinated equilibrium corresponds to the wage and participation threshold that would be selected if firms could coordinate on the monopsony's choice. To the southwest a continuum of additional uncoordinated equilibria are supported by agents beliefs about each other's strategies.

Figure 5 illustrates the best response correspondence of workers (dash) and best response function firms (dot-dash). These are mutual best responses on the 45 degree line from (r^L, w^L) to (r^C, w^C) and on the horizontal line (r^C, w^C) to $(\bar{b}, w^C)$.[14] Note that in the simple model expectations are formed at random.[15] In any equilibrium, these randomly formed expectations must be self-confirming: $w^* = w_0$ and $r^* = r_0$. To see how this works, consider the following: suppose the worker expects wages to be low. Her best response is to set a low participation threshold. The firm's best response is then to confirm the expectation by setting a low wage. Now consider the inverse of the problem: the firm expects the participation threshold to be low. Its best response is to set a low wage. The worker's best response is then to confirm the expectation by setting a low threshold. Thus, the expectations that support the equilibrium are self-confirming.

We categorize equilibria into two sets: coordinated and uncoordinated equilibria. We

[14]The appendix contains a diagram of the more complicated case that arises when the solution to the coordinated firms' problem is not unique.

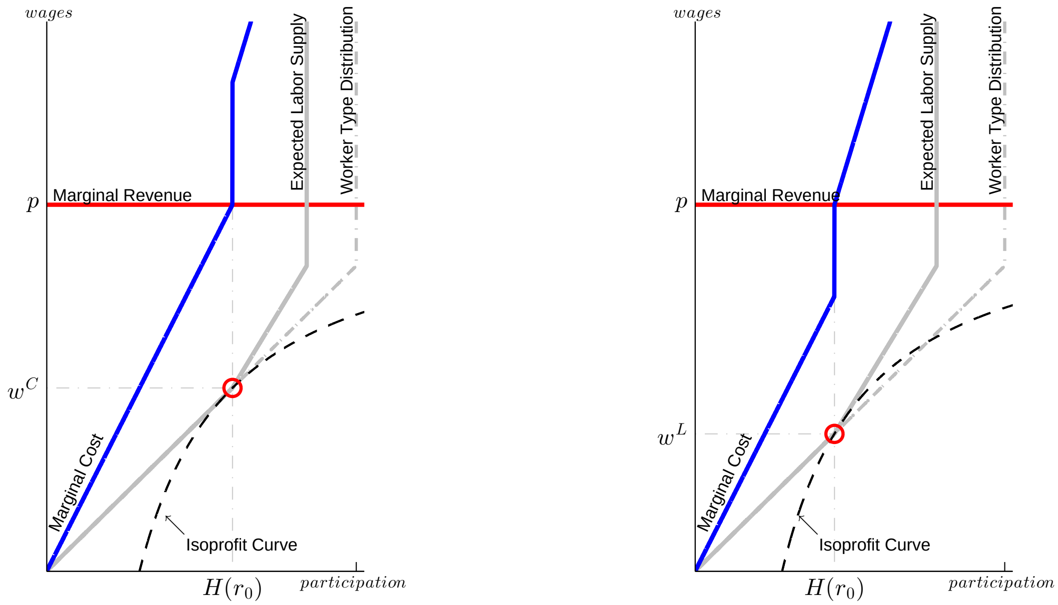
[15]In the stochastic economy we will consider how expectations come to be formed.

have considered expectations on the posted wage, w_0, and the participation threshold, r_0, that are formed in some sense randomly. Now consider the following sense of coordination: augment the game to allow the worker and the firm to meet prior to playing the game in order to form their expectations. Given that our game has both positive spillovers and strategic complementarities there exists a subset of equilibria that welfare dominate. Thus, in the pre-meeting the worker and firm gain from coordinating their expectations on wage and participation pairs consistent with the dominating equilibria. We label these as coordinated equilibria. The remaining equilibria exist only as equilibria of the game in which the pre-meeting is not allowed and expectations are formed randomly and subsequently self-confirmed.[16]

Figure 6 depicts the firm's optimization problem at the endpoints of the interval of possible equilibria: (r^L, w^L) and (r^C, w^C). w^C is defined as the wage choice made by the firm when it is weakly unconstrained by worker's participation threshold. Thus $r^C = w^C$ is the lowest threshold under which the firm is not constrained. Note that for thresholds below r^L the best response of the firm is to post a wage in excess of r^L, thus the strategies are *not* mutual best responses and can not be part of a self-confirming rational expectations equilibrium. In addition to these equilibria that are constrained by coordination failure, an additional set of equilibria exist such that the firm is not constrained by the participation choice of workers–wages are equal to w^C–and participation is at least as large as r^C. These are all possible equilibria since the worker is indifferent between participation thresholds that are weakly greater than the wage level.

[16]Note that coordination requires only strategic complementarity–consider the game "which side of the road to drive on"–however with positive spill-over one can institute coordination by allowing the worker and firm to agree to share the surplus in the pre-meeting.

Figure 6: Boundaries $-w^L$ and w^C- of the Uncoordinated Wage Set.



Note: w^C is defined as the wage choice made by the firm when it is weakly unconstrained by worker's participation threshold. Thus $r^C = w^C$ is the lowest threshold under which the firm is not constrained. r^L is defined as the lowest participation threshold such that this participation threshold and the wage choice that is the best response of the firm are mutual best responses.

4 A Two Sided Frictional Labor Market

We embed this stylized game in a frictional labor market that is endowed with appropriate features to justify the assumptions of our two player game–atomistic workers and atomistic firms meet via a process of sequential random matching–and show that firms' and workers' strategies in the multi-player game follow those of the two player game. Finally, we close the model with a standard matching function and the assumption of free entry into vacancy creation. Thus, equilibria of the full model are triplets: wage, participation threshold, and labor market tightness.

We posit that atomistic workers and atomistic firms meet via a process of sequential random matching and both discount the future at rate ρ. In the two-player game we assumed that the probability of a worker (firm) meeting a firm (worker) is exogenous to that player's strategy. In the full model this assumption is micro-founded by the assumption that every agent is atomistic. Thus, (1) the expectation of the strategy to be played by any given firm (worker) is equivalent to the average wage (participation threshold) and (2) the equilibrium job finding (filling) hazards are exogenous to each worker's (firm's) strategy. In the two-player game we assumed that the firm holds monopsony power and posts wages ex-ante. As Diamond (1971) shows, sequential random matching endows the firm with this monopsony power and the monopsony wage level prevails even in the limit as search friction fades.

We close the model by positing a standard matching function as in the Diamond-Mortensen-Pissarides (DMP) model. Firms post vacancies at flow cost, c and workers engage in search at zero cost. As in the baseline DMP model the flow of new matches is determined by the matching function, denoted as $m(U, V)$, where U is the mass of unemployed workers and V is the mass of vacancies.[17] Imposing Inada conditions and constant returns to scale, the job-finding rate of unemployed workers, $f(\theta) \equiv \frac{m}{U} = m(1, \theta)$, is increasing and concave in the market tightness defined as the ratio of vacancies to the unemployed, $\theta = \frac{V}{U}$. Analogously, the rate at which vacancies meet unemployed workers, $q(\theta) \equiv \frac{m}{V} = \frac{f(\theta)}{\theta}$, is a positive and decreasing function of market tightness.[18]

Note that since we have assumed random matching, the matching rate of workers with flow value of leisure above and below the equilibrium wage w^* is the same whenever they participate in the labor market.[19] Thus, the $i(1 - H(w^*))$ mass of workers who search for work but will reject wage offers of w^* are just as likely to meet a firm as the $H(w^*)$ mass of workers who search and will accept wage offer w^*. Note, some workers match but subsequently reject the wage offer. Thus, the job filling rate differs from the rate at which

[17] U and V are the mass of job seekers and vacancies. u and v are the exogenous odds of meeting a trading partner in the two player game. We will endogenize u and v here using the matching function.

[18] It is important to note here that our model differs in an important way from the classic model of upward sloping labor demand: Diamond (1982). In that model multiplicity derives from a thick market externality generated by increasing returns to scale in the matching function. In our model multiplicity derives from a pair of externalities, a thick market externality derived from workers' participation decision and a pecuniary externality derived from firms' wage posting decision. We follow the main stream DMP literature and impose constant returns to scale on our matching function. We appeal to the empirical results summarized in Pissarides and Petrongolo (2001) to justify this assumption.

[19] Note, under the assumptions defining the frictional labor market, the equilibrium wage coincides with the equilibrium wage in the two-player game.

vacancies meet employees. Since the $i(1 - H(w^*))$ mass of workers reject the wage offer whenever they meet a firm the vacancy filling rate is only:

$$\frac{H(w^*)}{H(w^*) + i(1 - H(w^*))} q(\theta) = \Lambda(w^*) q(\theta). \tag{9}$$

$\Lambda(w^*) \equiv \frac{H(w^*)}{H(w^*)+i(1-H(w^*))}$ is a measure of the severity of the congestion imposed by workers with leisure value above the equilibrium wage level searching for work while at the same time rejecting all offers. $\Lambda = 1$ occurs when workers with leisure value above the equilibrium wage never search and is the case when congestion is minimized. For $i > 0$ and $w^* < \bar{b}$, $\Lambda < 1$ and congestion drives a wedge between the rate at which vacancies meet employees and the rate at which jobs fill.

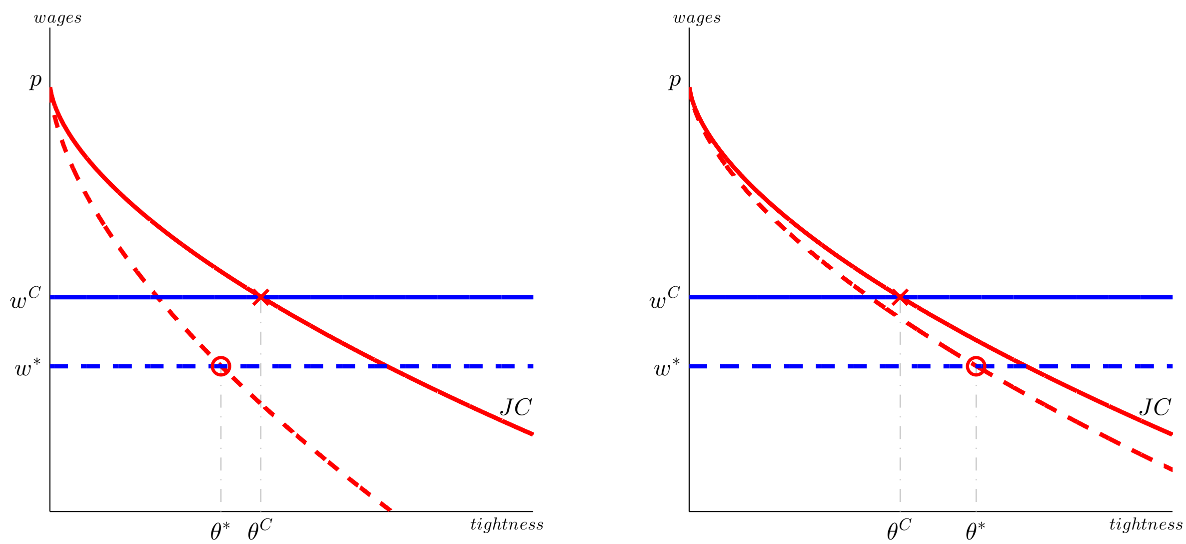
Job destruction is an exogenous shock that arrives with Poisson arrival δ. At the steady state, the flow into and the flow out of unemployment amongst workers with value of leisure below the participation threshold are equal, $(1 - u)\delta = f(\theta)u$, where u is, as in the two player game, the probability that a worker with value of leisure below the average wage level is unemployed. In addition some fraction of workers with value of leisure above the participation threshold also search; however, even when these workers meet firms they do not match. This gives the steady state mass of unemployed workers:

$$U = uH(w_0) + i(1 - H(w_0)) = \frac{\delta}{\delta + f(\theta)}(H(w_0) + i(1 - H(w_0))). \tag{10}$$

The appendix specifies the asset value equations for workers and firms. These yield identical participation and wage posting strategies as in the two-player game: the best response of workers is a participation threshold strategy as in equation (3) and the best response of firms is a wage posting strategy as in equation (8).

Given the wage, free entry into vacancy creation pins down the labor market tightness

Figure 7: Free Entry into Vacancy Creation.



as the value of a vacancy is driven to zero. Thus equilibrium tightness satisfies:

$$\frac{c}{q(\theta(w^*))\Lambda(w^*)} = \frac{p - w^*}{\rho + \delta}. \tag{11}$$

As is typical, the job creation condition is downward sloping while the wage schedule is weakly increasing (in our case flat), guaranteeing a unique labor market tightness for every aggregate wage level.

Definition 2. *A* self-confirming rational expectations equilibrium *of the frictional labor market is a tripple –* ***wage level, participation threshold, labor market tightness*** *– such that:*

1. *the wage level of each firm and participation threshold of each worker are mutual best responses.*

2. *labor market tightness satisfies the free entry condition, equation (11).*

Proposition 4. *A continuum of equilibria exists in the two-sided frictional labor market – wage levels in the interval* $[w^L, w^C]$ *– with higher welfare for higher wage levels.*

Proof. This is again a straightforward application of Cooper and John (1988) Proposition 5.

Proposition 2 enables us to rank equilibria in the frictional labor market in terms of the output gap defined as the difference between total output at an equilibria with an arbitrary wage level and total output at the constrained efficient equilibrium at wage level w^C.

Three comments are in order. First, the nature of the coordination failure is deeper in the multi-player economy. It is no longer sufficient for a single worker and single firm to meet and pre-arrange their expectations. It must be the case that each worker can coordinate her expectation with every firm and visa versa. This is because each faces a zero probability of meeting a specific worker or specific firm through matching in the frictional labor market. Instead we must think of a pre-meeting in which *all firms* can coordinate in order to provide a single wage signal on which workers can then coordinate their expectations of the wage level (and visa versa). In addition, note that the equilibrium prediction of the multi-player game is stronger than that of the two-player game in the sense that a unilateral deviation from the equilibrium by any player, even if observable to the other players, triggers no response since the odds of encountering the particular player through random matching are null. In the following sections we refer to an uncoordinated firm or uncoordinated worker as one who recognizes that her unilateral deviation from a given set of expectations does not impact the expectation of any other player.

Second, it is straightforward to observe that any minimum wage $\underline{w} \in (\underline{b}, w^C]$ is welfare improving. The minimum wage constrains the set of wages that may be consistent with rational expectations and thus ensures that higher welfare equilibria are picked on average. Setting $\underline{w} = w^C$ of course eliminates all the loss associated with coordination failure.

Third, a question remains of how firms and workers form expectations over the strategies of the other players in the market. In this baseline setup we can generate fluctuations in total output via a coordinated change in expectations: i.e. animal spirits. In this type of fluctuation agents form expectations on w_0 according to some aggregate signal unrelated to fundamentals–a sunspot–and best respond to the expected best response of other players. Shifting between a low wage and a high wage equilibrium will entail an increase in total output and total employment although there is no change in productivity per worker. The implication for unemployment remains ambiguous, however.

Figure 7 illustrates equilibrium labor market tightness conditional on a high wage level (depicted as the intersection of the solid lines in each panel) and on a low wage level (depicted as the intersection of the dashed lines in each panel) under two assumptions on the magnitude of congestion. Heavy congestion is illustrated in the left panel and light congestion in the right. Note that when both wage levels are weakly less than w^C the equilibria can be welfare ranked and we know that total output and total employment are larger in the higher wage equilibrium. The slope of the job creation condition depends on the participation threshold through the congestion effect. Fixing wages, more congestion implies a steeper job creation condition and therefore a looser labor market at any wage level. This is the difference between the dashed job creation conditions in the two panels. Fixing congestion, a lower wage implies a tighter labor market. If the increase in congestion triggered by moving from the high wage to the low wage equilibrium is large enough then unemployment will be higher in the low-output equilibrium. This is illustrated in the left panel. If the increase in congestion is mild unemployment *falls*. In this case the change in unemployment and the change in the output gap are in opposition. This is illustrated in the right panel.

Thus, without knowledge of the magnitude of the congestion effect the relation between unemployment fluctuations and fluctuations in output slack is ambiguous. In the following sections we focus on fluctuations driven by changes in a fundamental–labor productivity–rather than animal spirits alone. However, the ambiguity between unemployment changes and changes in the output gap will remain, enabling us to generate "job-lessness" early in recoveries and "wage-lessness" late. Which effect dominates is an empirical question. In Section 8.1 we find evidence that the congestion effect dominates, particularly after severe contractions.

5 Stochastic Economy

We now consider productivity as a stochastic process–we take p to be a martingale–and we consider how the history of productivity shocks might influence (expectations of) the aggregate wage level. Since workers may move freely between participation and non-participation and since free entry into vacancies guarantees that the value of a vacancy is zero in all periods, worker and firm value functions are independent of the shock process so long as shocks are not large enough to destroy existing matches. Further, we appeal to the short half-life of unemployment and vacancies to consider serial steady-states and abstract from dynamics in the following analysis (following the search literature, e.g. Mortensen and Nagypal (2007)).[20]

We note that our game is super-modular and impose the restriction that deviations from an existing wage-labor force participation pair occur only when unilateral deviation from the pair is the best response of some player. In other words, the equilibrium wage-labor force participation pair changes only when fundamentals evolve such that the existing pair is no longer in the set of possible equilibria. The new equilibrium is then selected from the new feasible set according to *best-response dynamics* as in Vives (1990, 2005); Cooper (1994).

We justify this refinement by noting that in such a situation a deviation from the initial existing wage and participation threshold pair would be the best response of an individual firm not only if no other firm deviates and workers' expectations do not change but also if all other firms' make the same deviation and workers expectations change accordingly. This follows from the supermodularity of the game.[21] Since all firms face the same problem, each should expect the others to make the same deviation. Also, since workers know all firms face the same problem they should expect each to make the same deviation. In other words, while

[20]Note that matches formed when wages are lower will be more robust to shocks, since the workers involved are selected on having lower flow value of leisure. This effect will rotate the job creation condition counter clockwise. This effect is outside the scope of this research and the results presented in the following are not qualitatively affected.

[21]See the Appendix.

any particular equilibrium may not be predictable, in the sense that many are possible and are contingent on expectations, deviations from one equilibrium to another *are* predictable and in this sense the new equilibrium can be learned by all players via the process of *dynamic best-response*.

5.1 Endogenous Rigidity in Wages and in Participation

Importantly, since there are a set of equilibria consistent with any given productivity level, transition between equilibria may require large shocks or an accumulation of small shocks. As a result aggregate variables, such as total output, may move sluggishly compared to the pace of labor productivity innovations.

Proposition 5. *For each existing wage level, w_0, and participation threshold, r_0, pair:*

- *There exists interval of productivity levels – (p^L, p^H) – for which no firm wishes to unilaterally deviate to a different wage:*

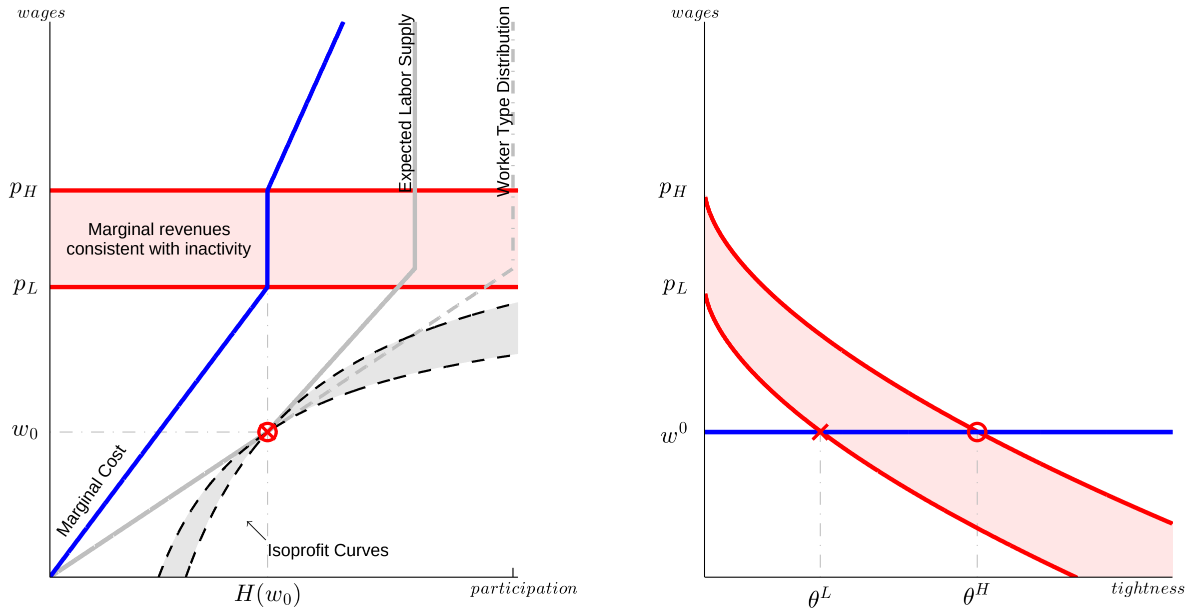
$$p^L = w_0 + \frac{H(w_0)}{h(w_0)} \qquad\qquad p^H = w_0 + \frac{H(w_0)}{ih(w_0)}.$$

- *There exists a corresponding interval of labor market tightnesses – (θ^L, θ^H) – where the lower bound solves the free entry condition, equation (11), for p^L and the upper bound for p^H.*

Proof. The proof follows from considering the first order conditions of the problem of a firm. This is essentially a corollary to the existence of the interval of equilibria on $[w^L, w^C]$ for any given productivity level.

For $i \in (0,1)$ this range of inactivity is depicted in Figure 8. The left panel presents the firm's monopsony problem constrained by the participation threshold as in Figure 4. For productivity in $[p^L, p^H]$ we see that the intersection between marginal cost and marginal revenue is at $H(w_0)$ and consistent with wage level w_0. The right panel presents the range of labor market tightness consistent with wage level w_0.

Figure 8: Inactivity Range.



5.2 Wage and Participation Revisions

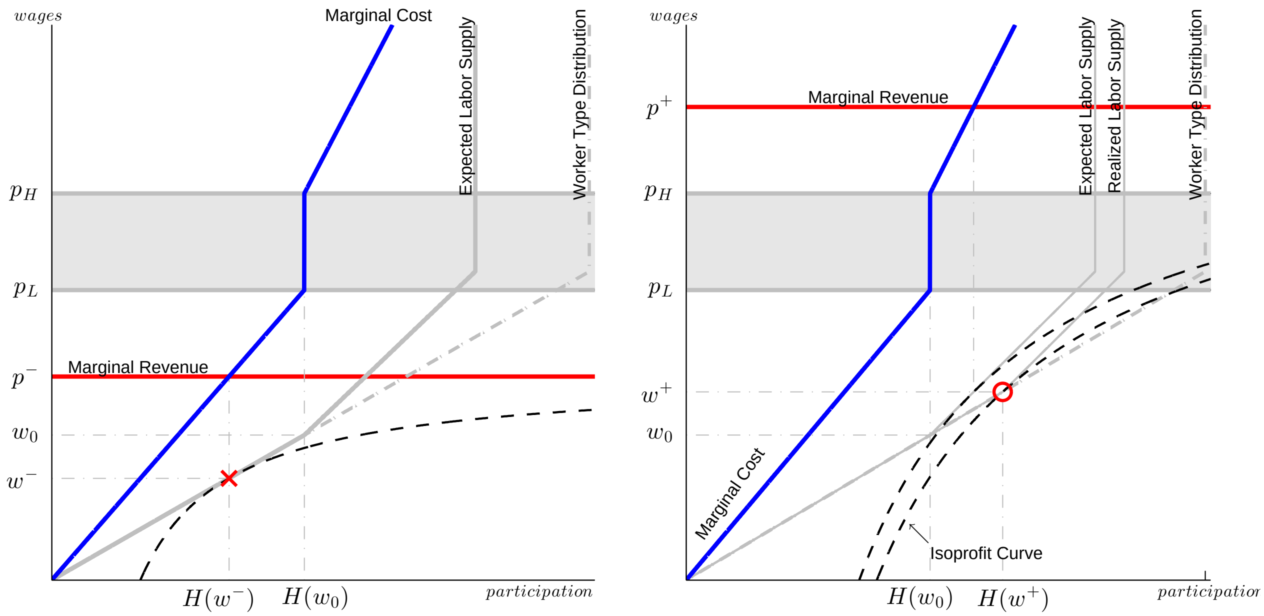
The model admits an innovation in average wages if a shock to productivity is sufficiently large, either rendering $p^- < p^L$ or $p^+ > p^H$. If a negative shock to productivity renders $p^- < p^L$, then, on impact, each firm is unconstrained by the existing participation threshold. Each selects a new wage such that $h(w^-)(p^- - w^-) = H(w^-)$. Since all firms choose the same unilateral deviation, w^- becomes the new aggregate wage level. This is illustrated in the left panel of Figure 9. In the left panel we see that after the drop in productivity each firm resets the target hiring level, $H(w^-)$, according to the intersection between marginal cost and marginal revenue.

In the second case a positive shock to productivity raises productivity to $p^+ > p^H$. Now each firm solves a constrained problem optimizing over mass of workers with flow value of leisure above existing wage level who search and are voluntarily unemployed at this existing wage level. Each firm unilaterally selects a new wage such that $ih(w^+)(p^+ - w^+) = H(w_0)(1 - i) + iH(w^+)$. This is illustrated in the right panel of Figure 9. In the left panel we see that after the increase in productivity each uncoordinated firm resets the target hiring level according to the intersection between marginal cost correspondence and marginal revenue. Note that this is to the right of the kink in the labor supply curve in the region where labor supply is inelastic. The new wage level is then found by evaluating this hiring level in the labor supply curve. Since all firms choose the same unilateral deviation, w^+ becomes the new aggregate wage level. Now, at this new aggregate wage level labor force participation will rise to $H(w^+)$, which is unexpectedly (from the firms perspective) higher than the targeted labor demand. Although each firm was unable to unilaterally implement this higher level of participation, the surplus participants make each firm better off and thus none renege on their wage deviation and due to the new inactivity range none wishes to deviate further (e.g. the game is supermodular).

5.3 Asymmetry

As the relative odds of meeting an inactive worker, i, rise to unity the upper bound, p^H, falls to the lower bound, p^L, and the unique equilibrium coincides with the equilibrium that would arise if workers and firms could coordinate on a wage and participation level for each level of labor productivity: a symmetric equilibrium. Also note as the relative odds of meeting an inactive worker fall to zero, the upper bound, p^H, rises without bound and the expanded model converges to a strictly asymmetric case in which the only possible revisions to wages and participation are downward. In the intermediate case we have a moderated degree of asymmetry: responses to positive and negative revisions to labor productivity *of the same magnitude* may lead to revisions in output of differing magnitude. We will call this hysteresis.

Figure 9: Revising the Equilibrium Wage.



5.4 Comments

Wage rigidity and the Barro (1977) critique

Our objective is to construct a Keynesian model in the sprit of the discussion of Howitt (1986). As such we strive for wage rigidity that is *evidence of* market failure while not directly causing market failure itself. The goal is to *end up* with endogenously rigid wages *because of* the market failure–as Keynes calls for–and *not* the other way around.

It is useful to dissect rigidity in our mechanism into ex-ante rigidity (before meeting any particular worker) and ex-post rigidity (after having met a particular worker). Firms in our model have ex-ante total wage flexibility; however, equilibrium posted wages are at times non-responsive to fundamentals due to coordination failure. Thus, ex-ante rigidity imposed by coordination failure is impervious to the Barro (1977) critique.

On the other hand, firms in our model experience ex-post total wage rigidity. We model wages this way, as does much of the wage posting literature, in order to capture an informational asymmetry between worker and firm: the firm does not know the worker's flow value of leisure. A typical critique of wage posting is that in the sub-game in which a rent-producing match is ruled out due to wage posting the informed party should reveal her type and bargaining should commence.[22]

In the limit as workers with flow value of leisure above the average wage never match, $i \searrow 0$, this suspect sub-game never occurs: only workers who will accept the posted wage search and all matches are consummated. Although ex-post wage rigidity never binds and ex-ante wages are flexible, wages are endogenously rigid on a vast region of values for the fundamental (labor productivity). For all realizations of productivity above p_L wages remain unchanged. Thus, for shocks that return p in this region wages are endogenously totally rigid. This case leads to the *largest* range of fundamentals for which wages are rigid in our model.

This result is too strong. Subjected to stochastic shocks, an economy in which workers with flow value of leisure above the average wage never search and never match will eventually converge to zero participation, zero employment, and zero output.

We take an intermediate value of i, allowing $i \in [0, 1]$, in order to admit the possibility of macroeconomic recovery following contractions. When we have $i > 0$ we are *partially* subject to Barro's criticism in the sense that ex-post some rent producing matches are ruled out. Of course, if negotiation is allowed in this sub-game, there will be cases when a mutually beneficial deal can be struck. Further, if such negotiation is permitted this will be known by workers and, as a result, beliefs consistent with inactivity will be difficult to support (at least at the threshold for participation predicted in the baseline model).

One possible solution to this sub-game that preserves our results is to assume that workers can trigger bargaining if they wish. If they do their bargaining power is zero. The result is that workers with $b < r_0 = w_0$ will never trigger bargaining. Workers with $b \in (w_0, p]$ may, and if they do they will receive wages equal to their flow value of leisure: so, they won't care if they do or don't bargain.

[22]This criticism is valid for any model of monopolistic pricing: given the flexibility the monopolist would always prefer to price discriminate and price discrimination is more efficient than monopoly pricing.

The Shimer (2005) puzzle: amplification versus a shifting Beveridge Curve

In constructing a region of indeterminacy regarding the wage level in equilibrium our model bears resemblance to a set of models that utilize wage rigidity to solve the Shimer (2005) puzzle: e.g. Hall (2005), Gertler and Trigari (2009), and Kennan (2010). Rigidity in our model can be used toward this objective; however, our main goal is the generation of hysteresis. With that in mind, models that focus strictly on amplification categorically fall short.

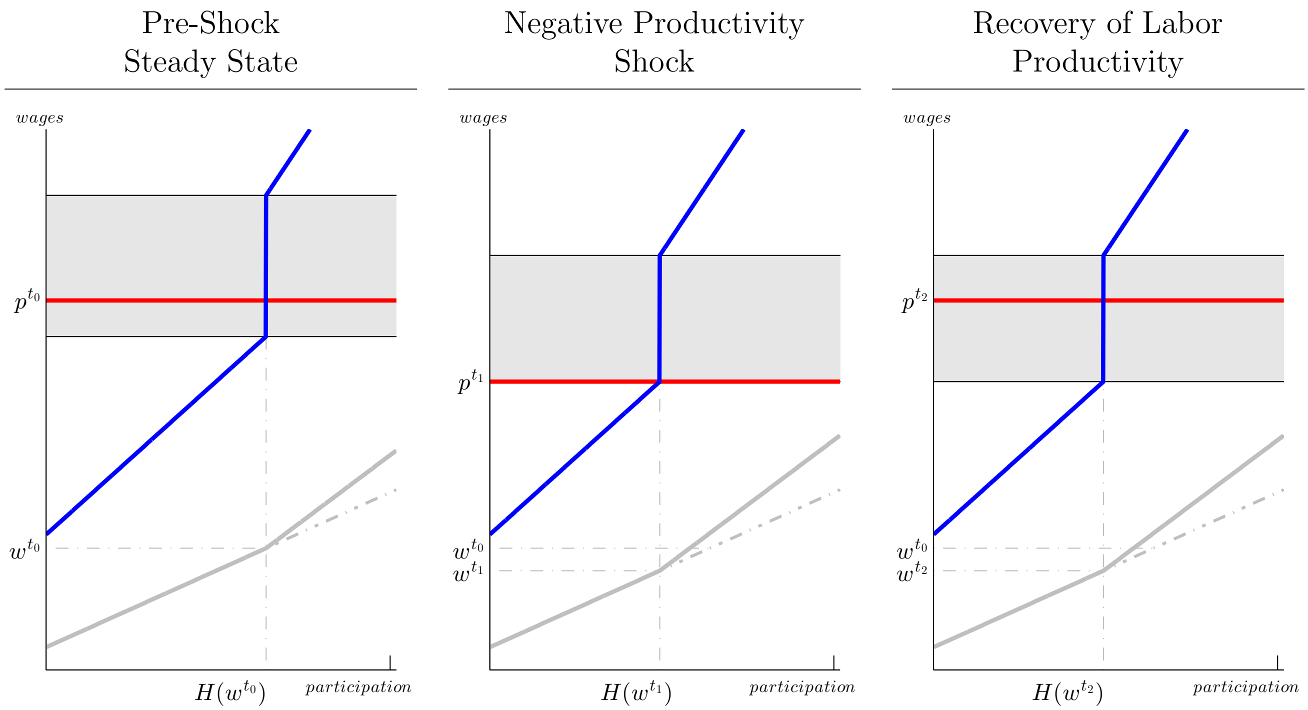
The logic is simple, in the absence of a congestion effect, a negative shock large enough to cause a downward revision in wages will lead to an *increase in amplification*. This increase is concentrated fully on the up-side: at the new lower wage the vacancy rate will be higher for any given level of productivity than it was at the old higher wage. The result is a "job-full" recovery. In contrast, the data shows that severe contractions are associated with outward shifts in the Beveridge Curve or "job-less" recoveries. Our model builds a congestion effect that is triggered at the same time as a downward revision of wages. Thus we are able to generate both amplification and a "job-less" recovery following severe contractions. We will show that, for us, amplification shows up in the "wage-less" phase of recovery.

6 Model Implications

6.1 Hysteresis

Asymmetry in wage and participation revisions gives the following stark implication: a negative shock to the marginal product of labor of magnitude X followed by the reversing positive but same sized shock of magnitude X *does not* necessarily imply that output and welfare return to the pre-shock levels. Whenever $p - X < p^L$ but $X < p^H - p^L$ this pair of shocks leads to a persistent reduction in welfare: hysteresis. In the post-shocks steady state participation, wages, total output, and employment will all fall short of the pre-shock steady

Figure 10: Hysteresis: Persistent Response to Transitory Negative Shock.



state.[23]

Figure 10 illustrates. In the left panel we illustrate a steady state at time t_0 in which productivity is within some inaction range. Wages and participation are at levels w^{t_0} and $H(w^{t_0})$ respectively. At time t_1 a shock arrives such that productivity at t_1 is outside and below the time t_0 inaction rage. As a result wages and participation fall to $w^{t_1} < w^{t_0}$ and $H(w^{t_1}) < H(w^{t_0})$ respectively. A new inaction range is established following the shock that partially over-laps and partially falls below the pre-shock inaction range. At time t_2 productivity recovers to the pre-shock level: $p^{t_2} = p^{t_0}$. However, even after the recovery productivity remains within the shock-state inaction range. As a result wages and participation *do not* respond to the recovery of productivity. Applying Proposition 4 we can welfare rank the time t_0 and time t_2 equilibria. In addition to wages and participation, total output and total welfare are lower at time t_2 than at time t_0. We have hysteresis.

[23]The implication for the unemployment rate is a bit more involved and depends on the relative magnitude of the wage rigidity and congestion effects. We treat this later in this section.

6.2 Inefficient Expansions

When $i > 0$, recovery from any given contraction is possible given a sufficiently positive sequence of shocks; however, the recovery of labor force participation, wages, and total output will still fall short of those that would be realized under coordination. Whenever some portion of workers do not to seek work, firms have insufficient incentive to raise wages in response to positive shocks. Thus, expansions, regardless of their length or speed, are insufficiently robust.

Proposition 6. *For $H(b)$ such that the upper support is not finite and for $i < 1$, employment and total output are fall short of constrained efficiency even after an arbitrarily long or steep expansion and there is a persistently open output gap.*

Proof. Recall that the wage set when productivity hits the upper bound consistent with a given wage level satisfies: $ih(w^+)(p^+ - w^+) = H(w_0) + i[H(w^+) - H(w_0)]$. For $i < 1$ we have $h(w^+)(p^+ - w^+) > H(w^+)$. So $w^+ < w^C(p^+)$ and the upward wage revision is inefficiently small. $\square$

6.3 "Job-less" and "Wage-less" Recoveries

Closing the model with free entry into vacancies, as in Section 4, the preceding two results can give rise to both "job-less" and "wage-less" recoveries. Job-lessness operates through the increase in congestion induced by the fall in the participation threshold during a severe contraction. Wage-lessness operates through an increase in the wage rigidity effect induced by the lower severe-contraction wage level. When the increase in congestion is not particularly severe, wage-lessness may dominate throughout a recession. However, whenever the drop in the participation threshold is severe enough to trigger a job-less recovery, job-lessness will occur during the early part of the recovery and will be followed by wage-lessness. Further, following a severe contraction, unemployment rates *must* for a time undershoot levels experienced at the same productivity level pre-contraction.

Figure 11: “Job-less” and “Wage-less” Recoveries.

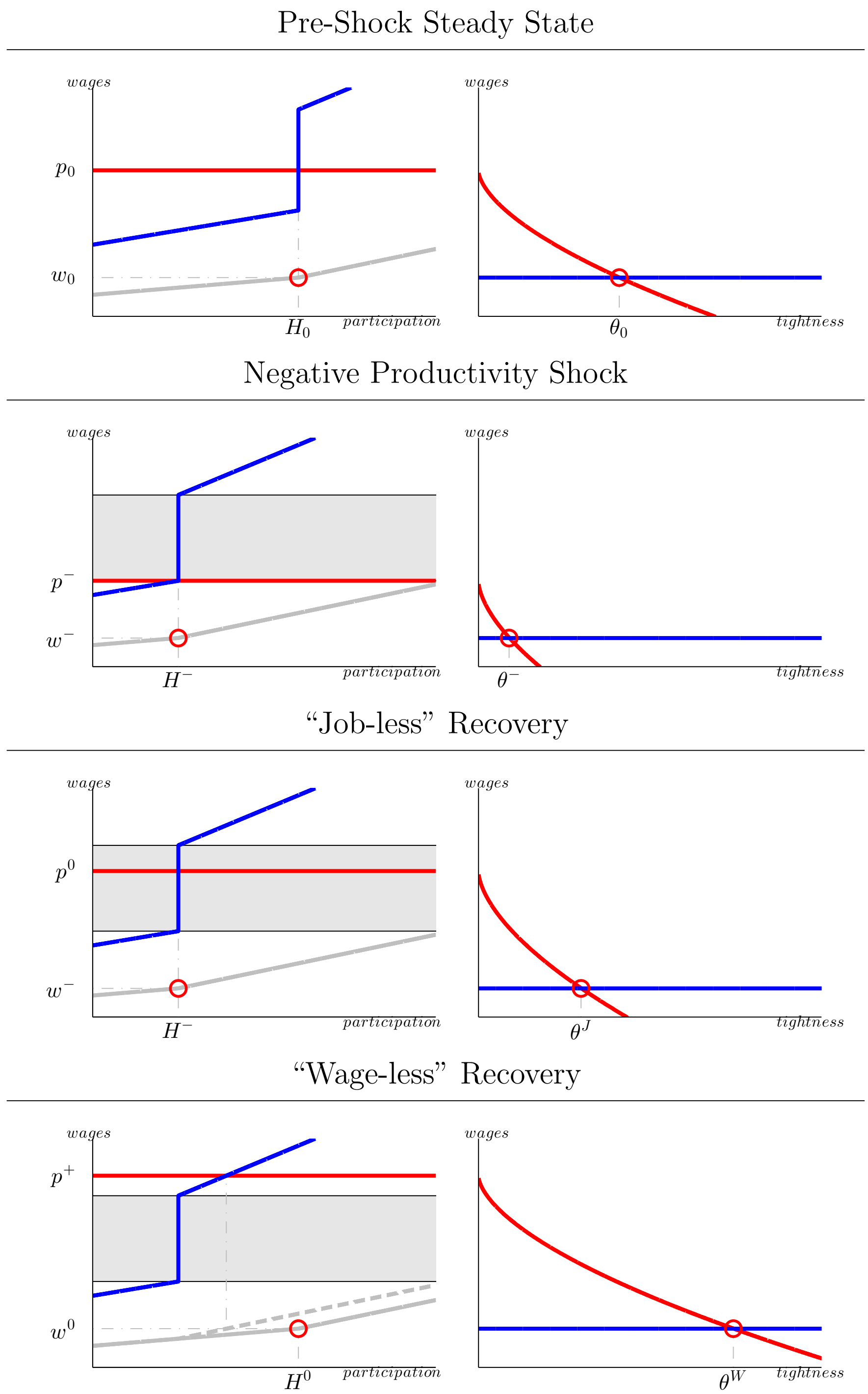


Figure 11 illustrates. In the top row we illustrate a steady state at time 0. At this pre-shock steady state, wages, participation, and labor market tightness are w^0, H^0, and θ^0 respectively. Our economy is then hit by a shock that lowers productivity to p^-, a value below the pre-shock inaction range. This is illustrated in the second row. As a result of the shock, wages and participation fall to w^- and H^- as firms lower the wage level due to their lower productivity (e.g. w^0 is now above the unconstrained wage choice in the shock state: $w^0 > w^C(p^-)$). Labor market tightness also falls. The drop in labor market tightness is the the result of three forces: the drop in productivity, an increase in congestion due to the drop in the participation threshold, and a drop in wages. The first two clearly push tightness down. That these dominate the third is a result of firms' optimization. Note that a fall in labor market tightness implies an increase in unemployment.

The third row illustrates wages, participation, and labor market tightness at a time when productivity has recovered to the pre-shock level. Since pre-shock productivity is in the inaction range determined in the shock state, wages and participation do not respond. As a result labor market tightness rises. Now tightness is a result of two of the three forces: the increased in congestion due to the lower participation threshold and the lower wage level. In our illustration the congestion effect dominates and tightness falls short of the pre-shock level. Note that a shortfall in tightness implies a higher level of unemployment: we have a job-less recovery.[24]

The final row illustrates the productivity shock required to return the economy to the pre-shock wage level following the contraction illustrated in the second row. Due to the asymmetry induced by the inaction range the required productivity level exceeds the pre-shock productivity level. Thus, labor market tightness in the period in which wages and participation recover *must* exceed the tightness experienced just before the impact of the first shock. Noting that congestion reduces as productivity approaches the level required to return wages to the pre-shock wage, we see that the wage rigidity effect must dominate the congestion effect before wages and participation recover to the pre-recession levels. Thus we have that at least the end phase of each recovery from a severe contraction must be wage-less. So, every job-less recovery must be followed by a wage-less recovery.

[24] The appendix illustrates the case when the increase in congestion is mild and a wage-less recover dominates throughout.

7 Policy

Pro-Cyclical Minimum Wages

Similar to the static model, a policy maker in the stochastic economy could improve the height of expansions and the robustness of a rebound in economic activity by implementing a pro-cyclical minimum wage policy. Any minimum wage policy satisfying $\underline{w}(p) \in (\underline{b}, w^C(p)]$ is welfare improving.

We offer a mechanism through which the policy maker can learn w^C. Suppose the policy maker can open a vacancy-call it public employment-and pay at a premium over the current average wage. Further, suppose that the existence of this vacancy and the wage paid is known to all workers. Further, suppose that the measure of the public sector vacancy is non-zero. If these conditions are satisfied then workers with flow value of leisure in excess of the current private sector wage but less than the public sector wage will strictly prefer to search for work and participation will rise. If this thickening of the labor market is efficient, private sector wages will rise to match the public sector wage as firms capitalize on the thick market. If private sector wages do not rise to match the public sector wage then the public sector wage is too high and induces an increase in congestion.[25]

Search-Contingent Unemployment Insurance

In the stochastic environment the policy maker has an additional policy lever to consider: search-contingent unemployment insurance. Suppose unemployment insurance is structured such that (1) payment is contingent on search and (2) insurance expires upon matching. Such a scheme does not alter the reservation wage policy of workers. However, the scheme breaks the indifference between inactivity and voluntary unemployment in favor of voluntary

[25]Note, the fact that we see private sector wages fall short of public sector wages; however, does not necessarily lead to the conclusion that the current equilibrium is efficient. It must also be the case that the public sector vacancies have non-zero mass. Thus, this is potentially instead an argument for pro-cyclical stimulus in the form of public sector hiring. This would increase the mass of the public sector vacancy during expansions and apply the needed signal of rising wages to induce labor market entry and allow the private sector to enjoy a thicker market.

unemployment for workers with high flow values of leisure. Recipients will search for employment with probability one when they would have otherwise searched only with probability i.

In the steady state such a policy clearly increases congestion and negatively impacts welfare. In the stochastic environment, however, there is a benefit. As i approaches unity, market thickness to the right and left of quantity of labor supplied at the participation threshold approaches each other. Thus the equilibrium wage-participation choice approach uniqueness at the coordinated equilibrium even in the face of stochastic shocks. The welfare gain from improving equilibrium selection must be balanced against the welfare loss from increased congestion in each steady state.[26]

8 Testing Some Implications if the Model in U.S. Data

8.1 Congestion and Match Efficiency

In Section 2 we highlighted evidence that contractions lead to persistent shifts in match efficiency from Diamond and Şahin (2015) and Barnichon and Figura (2015). Here we test the prediction of our model that these shifts are due to persistent shifts in labor force attachment triggered by severe contractions.

We begin by replicating the estimation of the matching function as in Barnichon and Figura (2015). We follow by assuming a constant returns to scale Cobb-Douglas matching function:

$$f = A\theta^{(1-\eta)}$$

[26]The current U.S. policy, under which search-contingent benefits expire a finite time after job-separation can exacerbate, rather than mitigate, losses due to labor force drop-out. To see this note that the longer a contraction persists then the fewer non-employed workers will qualify for search-contingent benefits. Thus, as a contraction continues the value of i should be falling. As a result, the positive shock needed to lift wages becomes larger as the contraction persists. Benefit extension during the Great Recession potentially extended the horizon during which the labor market is still relatively thick above the current wage level at the expense of increasing congestion over the same horizon. Ultimately, benefit extension is only effective if recovery occurs before (extended) benefits expire.

Table 3: Estimates of the Matching Function.

$$\ln f_t = \ln(A) + (1-\eta)\ln\theta_t + \varepsilon_t$$

	(1)	(2)
$\ln(A)$	-1.08***	0.37***
	(.008)	(0.080)
$(1-\eta)$	0.34***	0.26***
	(.010)	(0.009)
$\ln(male\ LFP)$		6.673***
		(0.366)
R^2	0.77	0.86
Sample (monthly frequency)	1967-2014	1967-2014

Source: Bureau of Labor Statistics, Current Population Survey; Conference Board, Help Wanted Index and Help Wanted Online Index; Bureau of Labor Statistics, Job Openings and Labor Turnover Survey. Vacancies are constructed from the Conference Board data as in Barnichon (2010) prior to 2001 and follow the Job Opening and Labor Turnover Survey thereafter. Male labor force participation is calculated from the Current Population Survey restricted to males between 25 and 55. Labor flows are calculated from the Current Population Survey monthly data using the matching procedure of Shimer (2012).

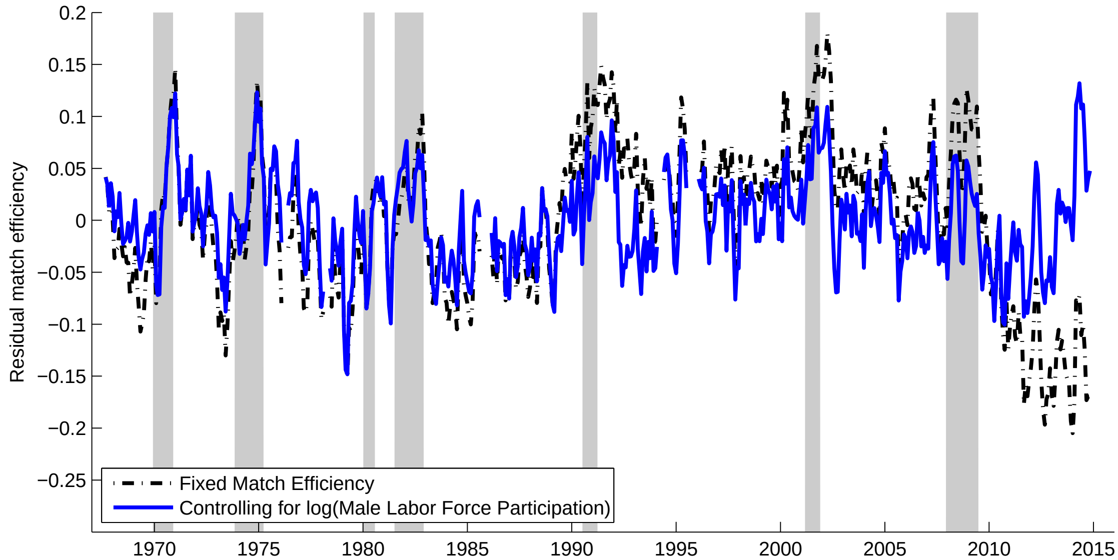
[a] Note: all regressions control for a linear time trend.

where η is the elasticity of the matching function with respect to the unemployment level. Log-linearizing we have:

$$\ln f_t = \ln(A) + (1-\eta)\ln\theta_t + \varepsilon_t$$

We estimate this regression using monthly data on unemployment and the job finding rate constructed from the Current Population Survey and monthly data on vacancies from the Conference Board and the Job Opening and Labor Turnover Survey (JOLTS). The job finding rate is constructed from labor flows calculated from the Current Population Survey monthly data using the matching procedure of Shimer (2012). The Conference Board data is adjusted to reflect the increased prevalence of online advertisement following Barnichon (2010). We normalize this series to the level of vacancies reported in JOLTS in January 2001 and use the JOLTS data following this date. The results of this regression are presented in

Figure 12: Match Efficiency.
(Sample Period: 1976–2015)



Source: Bureau of Labor Statistics, Current Population Survey; Conference Board, Help Wanted Index and Help Wanted Online Index; Bureau of Labor Statistics, Job Openings and Labor Turnover Survey. Vacancies are constructed from the Conference Board data as in Barnichon (2010) prior to 2001 and follow the Job Opening and Labor Turnover Survey thereafter. Male labor force participation is calculated from the Current Population Survey restricted to males between 25 and 55. Labor flows are calculated from the Current Population Survey monthly data using the matching procedure of Shimer (2012).

Table 3. A similar exercise instrumenting with lagged values yields similar results.

As in Barnichon and Figura (2015), residual reveals that it is predictable and thus the baseline regression is misspecified. We augment the regression, as our model suggests, to include time-varying match efficiency that depends on congestion, Λ:

$$\ln f_t = \ln(A) + \ln(\Lambda_t) + (1-\eta)\ln\theta_t + \varepsilon_t.$$

In addition to predictability of the errors, our model suggests that omission of Λ_t leads to omitted variable bias in the estimate of η. The job finding rate and Λ are obviously positively correlated (recall that lambda ranges between 0 and 1 with 1 representing zero congestion). The relationship between tightness and Λ is ambiguous, however. If the congestion effect dominates then a drop in Λ implies a drop in tightness. However, if the wage rigidity effect dominates then a drop in Λ implies an *increase* in tightness through the associated wage drop.

We measure Λ_t as the labor force participation of males with 1-30 years of potential experience who are civilians and not retired, ill, or disabled. We take this as a proxy for the labor force participation threshold noting that such males face no significant secular changes in participation incentives over this horizon and that participation is measured more accurately than wages since it is robust to compositor effects.

Our results indicate that male labor force participation is positively and statistically significantly related to the job finding rate. The elasticity of the job finding rate with respect to the male labor force participation rate is 6.7 percent. This suggests that if participation had not eroded over the observation period the job finding rate in 2009, following the Great Recession, would have been 0.29 on average as opposed to 0.18 as realized in the data.

The R^2 of our regressions jumps from 77 percent in the baseline regression to 86 percent when including male labor force participation as a measure of Λ_t. Further, Figure 12, illustrates that with the inclusion of this measure the residual is substantially less predictable. Thus, we conclude that the evidence with respect to match efficiency is in line with the prediction of our model: reduction in labor force attachment erodes match efficiency. Further, our regressions reveal a downward bias in the estimate of η when proxies for Λ are omitted. This tends to support that the congestion effect dominates the wage rigidity effect.

9 Conclusion

We contribute a theoretical model of hysteresis in the labor market to complement a substantial body of economic literature that shows that high unemployment, falling wages, and reduced economic activity can have lasting consequences. We model hysteresis as resulting from a coordination failure among atomistic firms and atomistic workers in a frictional labor market that features random search, *ex-ante* wage commitments, and the possibility of worker non-participation. This coordination failure results in a continuum of possible (fragile) equilibria with high-wage, high-participation equilibria welfare dominating low-wage, low-participation equilibria. We then introduce changes in labor productivity and show that if the transition among equilibria is governed by dynamic best-response then endogenous rigidity in wages and participation arises in response to transitory movements in labor productivity. The model is consistent with a host of other medium-run and business cycle facts following a severe recession: outward shifts in the Beveridge curve, job-less and wage-less recoveries, and a reduction in the sensitivity of wages to unemployment. Furthermore, expansions are insufficiently robust and in this sense recessions are scarring. This scarring effect can be mitigated by a state-dependent minimum wage policy which facilitates coordination on the efficient equilibrium.

References

Acemoglu, D. and R. Shimer (1999, November). Holdups and Efficiency with Search Frictions. *International Economic Review 40*(4), 827–49.

Albrecht, J., F. Postel-Vinay, and S. Vroman (2013, 05). An Equilibrium Search Model Of Synchronized Sales. *International Economic Review 54*(2), 473–493.

Barnichon, R. (2010). Building a composite help-wanted index. *Economics Letters 109*(3), 175 – 178.

Barnichon, R. and A. Figura (2015). Labor market heterogeneity and the aggregate matching function. *American Economic Journal: Macroeconomics 7*(4), 222–49.

Barro, R. J. (1977, July). Long-term contracting, sticky prices, and monetary policy. *Journal of Monetary Economics 3*(3), 305–316.

Beaudry, P., D. Galizia, and F. Portier (2015, June). Reviving the limit cycle view of macroeconomic fluctuations. Working Paper 21241, National Bureau of Economic Research.

Blanchard, O. J. and L. H. Summers (1986a, March). Hysteresis and the European Unemployment Problem. In *NBER Macroeconomics Annual 1986, Volume 1*, NBER Chapters, pp. 15–90. National Bureau of Economic Research, Inc.

Blanchard, O. J. and L. H. Summers (1986b, June). Hysteresis and the european unemployment problem. Working Paper 1950, National Bureau of Economic Research.

Blanchard, O. J. and L. H. Summers (1987). Hysteresis in unemployment. *European Economic Review 31*(1-2), 288–295.

Cooper, R. (1994). Equilibrium Selection in Imperfectly Competitive Economies with Multiple Equilibria. *The Economic Journal 104*(426), 1106–1122.

Cooper, R. and A. John (1988). Coordinating Coordination Failures in Keynesian Models. *The Quarterly Journal of Economics 103*(3), 441–463.

Diamond, P. (1982). Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy 90*(5), 881–94.

Diamond, P. and D. Fudenberg (1989, June). Rational Expectations Business Cycles in Search Equilibrium. *Journal of Political Economy 97*(3), 606–19.

Diamond, P. A. (1971, June). A Model of Price Adjustment. *Journal of Economic Theory 3*(2), 156–68.

Diamond, P. A. and A. Şahin (2015). Shifts in the Beveridge Curve. *Research in Economics 69*, 18–25.

Eeckhout, J. and I. Lindenlaub (2015). Unemployment cycles. IFS Working Papers W15/26, Institute for Fiscal Studies.

Elsby, M., B. Hobijn, and A. Şahin (2013). The decline of the u.s. labor share. *Brookings Papers on Economic Activity 47*(2 (Fall)), 1–63.

Elsby, M. W. L. and M. D. Shapiro (2012, June). Why Does Trend Growth Affect Equilibrium Employment? A New Explanation of an Old Puzzle. *American Economic Review 102*(4), 1378–1413.

Gertler, M. and A. Trigari (2009, 02). Unemployment Fluctuations with Staggered Nash Wage Bargaining. *Journal of Political Economy 117*(1), 38–86.

Golosov, M. and G. Menzio (2015, November). Agency Business Cycles. NBER Working Papers 21743, National Bureau of Economic Research, Inc.

Hall, R. E. (2005, March). Employment Fluctuations with Equilibrium Wage Stickiness. *American Economic Review 95*(1), 50–65.

Howitt, P. (1986). The Keynesian Recovery. *The Canadian Journal of Economics / Revue canadienne d'Economique 19*(4), 626–641.

Juhn, C., K. M. Murphy, and R. H. Topel (2002). Current Unemployment, Historically Contemplated. *Brookings Papers on Economic Activity 2002*(1), 79–116.

Juhn, C., K. M. Murphy, R. H. Topel, J. L. Yellen, and M. N. Baily (1991). Why has the Natural Rate of Unemployment Increased over Time? *Brookings Papers on Economic Activity 1991*(2), 75–142.

Kaplan, G. and G. Menzio (2013, February). Shopping Externalities and Self-Fulfilling Unemployment Fluctuations. NBER Working Papers 18777, National Bureau of Economic Research, Inc.

Karabarbounis, L. and B. Neiman (2014). The global decline of the labor share. *The Quarterly Journal of Economics 129*(1), 61–103.

Kennan, J. (2010). Private information, wage bargaining and employment fluctuations. *The Review of Economic Studies 77*(2), 633–664.

Krugman, P. (1991). History Versus Expectations. *The Quarterly Journal of Economics 106*(2), 651–667.

Morris, S. and H. S. Shin (2000, September). Global Games: Theory and Applications. Cowles Foundation Discussion Papers 1275R, Cowles Foundation for Research in Economics, Yale University.

Mortensen, D. and E. Nagypal (2007, July). More on Unemployment and Vacancy Fluctuations. *Review of Economic Dynamics 10*(3), 327–347.

Murphy, K. M. and R. Topel (1997). Unemployment and Nonemployment. *The American Economic Review 87*(2), 295–300.

Pissarides, C. A. and B. Petrongolo (2001, June). Looking into the black box: A survey of the matching function. *Journal of Economic Literature 39*(2), 390–431.

Schaal, E. and M. Taschereau-Dumouchel (2015). Coordinating Business Cycles. 2015 Meeting Papers 178, Society for Economic Dynamics.

Shimer, R. (2005, March). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *The American Economic Review 95*(1), 25–49.

Shimer, R. (2012, April). Reassessing the Ins and Outs of Unemployment. *Review of Economic Dynamics 15*(2), 127–148.

Summers, L. H. (1992, October). *Understanding Unemployment*. The MIT Press.

Vives, X. (1990). Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics 19*(3), 305 – 321.

Vives, X. (2005). Complementarities and Games: New Developments. *Journal of Economic Literature 43*(2), 437–479.

A Appendix

We can write out the asset value equations faced by workers and firms. For workers we have:

$$\begin{aligned}
\rho\mathcal{I}(b,w_0) &= b \\
\rho\mathcal{U}(b,w_0) &= b + f(\theta)max\{0,[\mathcal{W}(b,w_0)-\mathcal{U}(b,w_0)]\} \\
\rho\mathcal{W}(b,w_0) &= w_0 + \delta[\mathcal{U}(b,w_0)-\mathcal{W}(b,w_0)]
\end{aligned}$$

The first, $\mathcal{I}(b,w_0)$, captures the asset equation for inactive workers.[27] The second, $\mathcal{U}(b,w_0)$, =, captures the asset equation for searching workers, these receive flow value b and at hazard $f(\theta)$ receive job offers which yield option value $max\{0,[\mathcal{W}(b,w_0)-\mathcal{U}(b,w_0)]\}$. The third, $\mathcal{W}(b,w_0)$ captures the asset equation for employed workers, these receive flow value w_0 and at hazard δ are separated to unemployment.

For firms we have:

$$\begin{aligned}
\rho\mathcal{V}(w,r_0) &= -c + q(\theta)\Big[\mathbb{I}_{\{w\le b_0\}}\tfrac{uH(w)}{uH(r_0)+i(1-H(r_0))} \\
&\quad +(1-\mathbb{I}_{\{w\le b_0\}})\tfrac{uH(r_0)+i[H(w)-H(r_0)]}{uH(r_0)+i(1-H(r_0))}\Big][\mathcal{J}(w,r_0)-\mathcal{V}(w,r_0)] \\
\rho\mathcal{J}(w,r_0) &= p - w + \delta[\mathcal{V}(w,r_0)-\mathcal{J}(w,r_0)]
\end{aligned}$$

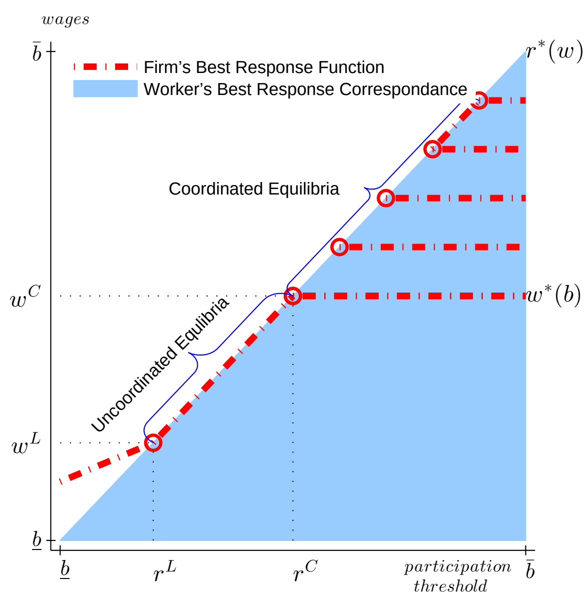
The first asset equation, $\mathcal{V}(w,r_0)$, captures the value of a open vacancy. The vacancy costs the firm a flow of c and with hazard $q(\theta)$ the firm meets and makes the posted wage offer w_0 to a worker. The bracketed term follows the logic of the firm's problem in the two player game: higher wage offers are accepted by a larger fraction of the workers that the firm might meet. The final term is the option value of forming a match. The second asset equation, $\mathcal{J}(w,r_0)$ captures the value of a filled job. This is the flow rent $(p-w)$ and the hazard of separation times the option value of separation.

[27] Note that in steady state this is independent of the wage level. When adding shocks we note that as long as there is no barrier to reentering the unemployment pool following a shock the independence is preserved.

We posit free entry into vacancy creation. This drives the value of a vacancy to zero. From these assumptions it is straightforward to show that the participation threshold (wage) strategies played by workers (firms) in this frictional market are equivalent to the strategies played by the worker (firm) modeled in the two-player game. Note that this implies that $(1-v) = \frac{q(\theta)}{\rho+\delta}$.

A.1 Non-uniqueness in the coordinated firms' problem

Figure 13: Best response correspondence with non-unique solution to the monopsony problem.



Note: Multiple coordinated equilibria correspond to the wage and participation threshold that would be selected if firms could coordinate on the monopsony's choice. To the southwest of the smallest of these is a continuum of additional uncoordinated equilibria are supported by agents beliefs about each other's strategies.

Multiple coordinated equilibria may exist if the monopsony's problem has multiple solutions. In the region above the smallest wage that solves the monopsony's problem, $\underline{w}^C$, all equilibria are efficient. To see this, note that should workers expectations fall short of what is required to maintain one of these equilibria, then $\underline{w}^C$ is always a possible best response of the firm. Additionally, any solution to the monopsony's problem that lies below workers' expected wage level is also a best response.