Learning When to Say No*

David Evans
University of Oregon

George W. Evans
University of Oregon and University of St Andrews

Bruce McGough
University of Oregon

February 20, 2019

Abstract

We consider boundedly-rational agents in McCall's model of intertemporal job search. Agents update over time their perception of the value of waiting for an additional job offer using value-function learning. Using a first-principles argument we show asymptotic convergence to fully optimal decision-making. We study transitional learning dynamics using simulations. Structural change induces two important qualitative features. First, an increase in benefits or the median wage causes a dramatic spike in the unemployment rate under rational expectations that is attenuated or nonexistent under learning. Second, a decrease in the median wage causes significant overshooting of the unemployment rate for boundedly-optimal agents.

JEL Classifications: D83; D84; E24

Key Words: Search and unemployment; Learning; Dynamic optimization; Bounded rationality.

1 Introduction

We reconsider the partial-equilibrium labor search model due to McCall (1970) in which a worker must decide whether to work at a given wage or to wait and search

*Financial support from National Science Foundation Grant No. SES-1559209 is gratefully acknowledged.

for a better wage. The behavior of a fully rational agent in the setting is well understood; however, this behavior requires considerable knowledge and sophistication. To address this issue, and thus add additional realism to the model, we adopt a bounded optimality approach along the lines of Evans and McGough (2018a) in which agents make decisions based on perceived trade-offs. The perceptions of these trade-offs are revised as experience is gained and new data become available.

We begin by developing our version of the McCall model, based on the presentation in Ljungqvist and Sargent (2012),[1] which obtains a solution under rational expectations (RE) and optimal decision-making. Then, using value-function learning as in Evans and McGough (2018a), we develop a framework for boundedly rational decision-making. Under very general conditions, we show directly, using the martingale convergence theorem, that asymptotic fully optimal decision-making obtains. Through numerical simulations, we study transitional dynamics under learning. These dynamics are distinct from their RE counterpart, and would plausibly arise when there are changes in policy or structure.

A key feature of our approach is that in making their decisions workers incorporate several structural features of the economy that they know, while learning over time about a key but unknown sufficient statistic for optimal decision-making. This unknown sufficient statistic, which we denote by Q^*, measures the agent's expected discounted utility when they are unemployed and waiting for a random wage offer. As is well-known, optimal decision-making in this setting is characterized by a reservation wage w^* that is pinned down by Q^*. Under boundedly rational decision-making with adaptive learning, agents use an estimate Q of Q^* to make decisions given their knowledge of the unemployment benefit level b and the probability α per period of job separation when employed. Their estimate Q determines their corresponding reservation wage $\bar{w}$, and thus their boundedly optimal decisions.

The estimate Q of Q^* is updated over time based on observed wage offers. It is natural to assume that both unemployed and employed workers observe a (possibly small) sample of wage offers; agent's update their estimate Q based on this sample. Our central theoretical result is that this procedure asymptotically yields fully optimal decision making: over time agents learn Q^*. We emphasize two distinct features of our result that are particularly attractive. First, agents do not need to have any knowledge of the distribution of wages; and second, their computations are simple as well as natural: they do not need to iterate a value function, or even to know of the existence of Bellman's equation.

After establishing these results we turn first to the comparative statics and dynamics under full optimality arising from changes in the unemployment benefit level or the median of the wage distribution. We decompose the RE comparative statics into the direct effects, i.e. those effects induced by changes in structure holding fixed beliefs, and into the indirect effects resulting from the changes in beliefs induced by

[1] Ljungqvist and Sargent (2012) use a linear specification of utility and focus solely on the rational expectations solution.

the structural change. Of particular interest is the "hazard rate" h that measures the probability of an unemployed agent accepting (and keeping) a job in a given period, together with its reciprocal δ, which measures the expected duration of a newly unemployed worker.

In response to a rise in b, the hazard rate falls and the expected duration rises: this reflects the corresponding increase in Q^* and that both direct and indirect effects are negative. On the other hand, in the presence of an increase in the median wage, the direct and indirect effects compete: for fixed beliefs, new wage offers are more likely to be viewed as attractive; however, because Q^* rises in response to the increase in μ, any particular offer is less appealing. For this reason, a condition is required in order to guarantee that the hazard rate rises and expected duration falls in response to a rise in the median wage – an outcome we view as natural and which we adopt as an assumption for our numerical work.

In our partial-equilibrium setting with a continuum of rational agents, the economy has a unique stationary distribution, and associated to this distribution is a natural measure of the unemployment rate, which, as we show, is inversely related to the hazard rate. Thus, an increase in benefits leads to an increase in steady-state unemployment, and an increase in the median wage leads to a fall in steady-state unemployment under our stated conditions. In response to a structural shock, the unemployment rate even under RE does not immediately jump to its new steady state level; in fact there are sometimes important transitional dynamics. For example, on impact an increase in benefits leads to a dramatic rise in unemployment followed by monotonic convergence to the new steady state. The intuition for this outcome is straightforward: the rise in benefits raises the reservation wage, which induces quits by employed workers with now inadequate wages. In contrast, a fall in benefits results in no similar spike. The response to a change in median wage is more nuanced, and discussed in detail in Section 4.3.

Under RE both the direct and indirect effects on w^* take place simultaneously; however, under adaptive learning only the direct effects on $\bar{w}$ are realized instantaneously, with the indirect effects materializing gradually through the adjustment of Q in response to new data. To examine this slow emergence of indirect effects, we turn to comparative dynamics analysis based on simulations. We parameterize our model and simulate the collective behavior of 6000 boundedly rational agents, and compare the outcomes to the collective behavior of 6000 rational agents. The impact effect of an increase in benefits on the unemployment rate under learning is greatly muted by the absence of the indirect effect: a smaller upward spike is observed and there is rapid convergence to the new steady state. Since learning agents do not observe or know the wage distribution, there is no direct effect on their behavior of an increase in the median wage, and therefore, in contrast to the RE case, no corresponding spike in the unemployment rate; instead, the unemployment rate converges quickly to its new steady state level. On the other hand, a fall in the median wage under learning leads to a spike in unemployment that overshoots its new higher steady-state level:

intuitively, learning agents are unaware that the wage distribution has deteriorated and therefore in the short run reject offers that ultimately they would accept.

The approach presented in this paper is related to several approaches in the literature that model boundedly rational decision-making. Like the adaptive least-squares learning approach in macroeconomics, e.g. Bray and Savin (1986), Marcet and Sargent (1989) and Evans and Honkapohja (2001), which focuses on least-squares learning, we consider decision-making procedures that, while not fully rational, have the potential to converge to rational expectations and fully optimal decision-making over time. Like Marimon, McGrattan and Sargent (1990), Preston (2005) and Cogley and Sargent (2008), our framework has long-lived agents that must solve a challenging dynamic stochastic optimization problem. In these settings two issues are of concern: (i) there are parameters that govern the state dynamics that may not be known; and (ii) the assumption that agents know how to solve dynamic stochastic programming problems is implausibly strong.

Cogley and Sargent (2008) examine the first issue carefully in the context of a permanent-income model with risk aversion. In their setting income is assumed to follow a two-state Markov process with unknown transition probabilities, which takes it outside the usual dynamic programing framework, and they consider two alternative approaches to decision making. The first is to treat their agents as Bayesian decision-makers following a fully optimal decision rule within an expanded state space in which the programming problem has a time-invariant transition law. This requires considerable sophistication and expertise for the agent as well as a finite planning horizon to make the problem tractable. The second approach is to employ the "anticipated utility" model of Kreps (1998), in which agents make decisions each period by solving their dynamic programing problem going forward based on current estimates of the transition probabilities. This procedure is boundedly rational in the sense that agents ignore the fact that their estimates will be revised in the future, but is computationally simpler; in their set-up Cogley and Sargent found the fully optimal procedure provided only a small improvement in decision-making.

Bayesian decision-making has also been used in boundedly-rational settings. Adam, Marcet and Beutel (2017) have shown how to implement this approach in an asset-pricing environment. In their set-up agents are "internally rational," in the sense that they have a *prior* over variables exogenous to their decision-making that they update over time using Bayes Law, though these beliefs may not be externally rational in the sense of fully agreeing with the actual law of motion for these variables. By imposing appropriate simple natural forms for these beliefs, it is possible to solve for the corresponding solution to avoid the expanded state-space issues encountered by Cogley and Sargent (2008) and solve the agents' dynamic programing problem.

The anticipated utility framework has also been employed in adaptive learning set-ups in which long-lived agents use least-squares learning. Preston (2005) developed an approach in which agents estimate and update over time the forecasting models for relevant variables exogenous to the agents decision-making. For given forecasts of these variables over an infinite horizon, agents make decisions based on the solution to their dynamic optimization problem. Again, these decisions are boundedly optimal in the sense that the procedure does not take account of the fact that their estimates of the parameters will change over time.

The approach adopted in the current paper is closest to the general bounded-optimality framework of Evans and McGough (2018a). In their approach infinitely-lived agents optimize by solving a two-period problem in which a suitable variable in the second period encodes benefits for the entire future. Their primary focus is on shadow-price learning, in which the key second-period variables are shadow prices for the endogenous state vector, but they also show how a value function learning approach can equivalently be employed in a setting with continuously measured state variables. Evans and McGough (2018a) use the anticipated utility approach and obtain conditions under which an agent can learn to optimize over time. Evans and McGough (2018b) show how to apply shadow-price learning within a wide range of DSGE macroeconomic frameworks.

The current paper applies a version of value-function learning in a discrete choice setting in which a worker must choose whether or not to take a wage offer. In our McCall-type set-up, the single sufficient statistic needed is the value of the dynamic optimization problem when the agent is unemployed and facing a random wage draw. We show how, given an estimate of this value, an agent can make boundedly optimal decisions under the anticipated utility assumption, and we demonstrate that when agents in addition use a natural adaptive-learning scheme for updating their estimates over time, they will asymptotically learn with probability one how to make optimal decisions within a stationary environment. Our model of boundedly optimal decision-making also embeds naturally in a model populated by a large number of agents with an economy subject to structural change, enabling us to study transitional learning dynamics.

Our framework is also related to the "Q-learning" approach developed originally by Watkins (1989) and Watkins and Dayan (1992) as well its extensions to temporal difference learning from the computer science literature.[2] In that approach agents make decisions based on estimates of quality-action pairs, with the quality function updated over time. As in the current paper the Q-learning approach is motivated by the Bellman equation, but it is typically and most effectively implemented in set-ups in which the state as well as action spaces are finite. In our set-up agents must make decisions when facing a continuously-valued wage distribution, where the distribution is unknown to the agents; furthermore, when making their boundedly optimal choices, our agents are able to incorporate features of the transition dynamics that are known to the agents, including separation rates and benefit levels.

Our paper proceeds as follows. Section 2 outlines the environment. Section 3 presents our model of boundedly optimal decision-making. Section 4 contains our

[2]See Sutton and Barto (2011) for a detailed introduction to reinforcement learning and in particular temporal difference learning.

comparative statics analysis as well as numerical simulations of unexpected structural change. Section 5 concludes.

2 The Model and Optimal Decision-making

We consider an infinitely-lived agent who receives utility from consumption via the instantaneous utility function u. Time is discrete. At the beginning of a given period the agent receives a wage offer, and decides whether or not to accept it. The wage offer is drawn from a distribution that depends on whether the agent was employed or unemployed at the end of the previous period. If the agent was employed, her wage in the previous period constitutes her wage offer in the current period. If the agent was unemployed in the previous period, she receives a wage offer w drawn from a time-invariant exogenous distribution F (density dF). In either case, the agent must decide whether or not to accept the offer. Intuitively, a wage offer should be viewed as an option, purchased by the firm, to employ the agent. The purchase price of the option is the wage.

If the wage offer is not accepted the agent is unemployed in the current period, and receives an unemployment benefit $b > 0$; and, because she is unemployed at the end of the current period, she will receive a wage offer drawn from F at the beginning of the next period. If the offer is accepted then the agent receives the wage w in the current period. With probability $1 - \alpha$ the firm exercises the option to employ the agent, and, because she is employed at the end of the current period, she will receive the same wage offer in the next period. With probability α the firm does not exercise the option to employ the agent, and consequently she is unemployed in the current period; thus, in the next period, she will receive a wage offer drawn from F. We remark that, under full rationality, an agent employed in the preceding period will always accept her wage offer in the current period; however, under bounded rationality, previously employed agents may decide to enter unemployment as their understanding of the world evolves.

We make the following assumptions to ensure the that the worker's problem is well behaved, which we set out for future reference:

Assumption A:

1. u is twice continuously differentiable, with $u' > 0$ and $u'' \leq 0$.
2. F has support $[w_{\min}, w_{\max}]$, where $0 < w_{\min} < w_{\max}$.
3. All wage draws are independent over time and across agents.
4. $0 < \alpha < 1$.

The first two items ensure the existence of and continuity of the worker's value function, while the third item guarantees that the worker's optimal value of search does not depend on additional state variables.

It remains to specify how agents make decisions, i.e. whether or not to accept the wage offer. In this Section we adopt the conventional assumption that agents are fully rational and we characterize the corresponding optimal behavior. In Section 3.1 we instead model agents as boundedly rational, and Section 3.2 characterizes optimal behavior as a fixed point of an induced map on beliefs. Section 3.3 completes our model of boundedly optimal decision-making with an adaptive learning story of how agents might update their beliefs over time, and proves our central result that, under suitable assumptions, boundedly optimal decision-making converges to fully optimal behavior.

The fully optimal agent makes decisions by solving the following programming problem:

$$\begin{aligned} V^*(w_0) &= \max_{\{a_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c(a_t, w_t)) \\ & \quad w_{t+1} = g(w_t, a_t, \hat{w}_{t+1}, s_{t+1}). \end{aligned} \tag{1}$$

Here $a_t \in \{0,1\}$ is the control variable identifying whether the job is accepted ($a_t = 1$) or not ($a_t = 0$), w_t is the endogenous state variable corresponding to the wage offer in period t, $\hat{w}_{t+1}$ is an i.i.d. random variable drawn from F, and $s_{t+1} \in \{0,1\}$ is an i.i.d. random variable taking on the value 0 with probability α, capturing the probability with which a given firm chooses not to exercise its option to employ the agent. Finally, the functions c and g are given as follows:

$$c(a,w) = \begin{cases} w & \text{if} \quad a = 1 \\ b & \text{if} \quad a = 0 \end{cases} \quad \text{and } g(w,a,\hat{w},s) = \begin{cases} \hat{w} & \text{if} \quad a = 0 \text{ or if } a = 1 \text{ and } s = 0 \\ w & \text{if} \quad a = 1 \text{ and } s = 1 \end{cases}.$$

The associated Bellman functional equation may be written as

$$\begin{aligned} V^*(w) &= \max_{a \in \{0,1\}} u\left(c(a,w)\right) + \beta E\left(V^*(w')|a,w\right) \\ & \quad w' = g(w,a,\hat{w},s), \end{aligned} \tag{2}$$

with the expectation E taken over random variables $\hat{w}$ and s. We note that, because of the properties of u, the finite support of the distribution F, and the compact (finite) control space, the Principle of Optimality implies that the solution to the Bellman equation (2) corresponds to the value function associated with the sequence problem (1), which is why we can use V^* in both equations.

The optimal value of $V^*(w)$ of having a wage offer w in hand allows us to define

$$Q^* = E\left(V^*(\hat{w})\right) \equiv \int_{w_{\min}}^{w_{\max}} V^*(\hat{w}) dF(\hat{w})$$

and note that Q^* is the value, under optimal decision-making, associated with being unemployed at the start of the period before $\hat{w}$ is realized. Moreover, as we will see in our introduction of bounded optimality, Q^* encapsulates all of the complicated features of this problem: that the wage offer distribution may not be known and that, even conditional on knowing the wage offer distribution, making optimal decisions requires solving a complicated fixed point problem.

3 Boundedly Optimal Decision-making

In this Section we specify how boundedly optimal agents make decisions, which requires allowing for an explicit dependence of the value function on beliefs. First in section 3.1 we show how boundedly optimal decision-making can be formulated in terms of an agent's perception of the expected discounted utility of receiving a random wage draw, a value we denote by Q. We note that only unemployed agents receive random wage draws; thus, Q may be interpreted as the value associated with being unemployed. In section 3.2 we demonstrate that optimal behavior can be viewed as a special case, i.e. $Q = Q^*$. Finally, in section 3.3 we show that under a natural updating rule the agent's perceptions Q converge over time to Q^*, i.e. agents learn over time to make optimal decisions.

3.1 Decision-making under subjective beliefs

Denote by Q the agent's current perceived (i.e. subjective) value of a random wage offer drawn from F. Let $V(w, Q)$ denote the perceived value of a wage offer w. With this notation we assume that boundedly optimal agents with beliefs Q make decisions by solving the following optimization problem

$$V(w, Q) = \max \{u(b) + \beta Q, u(w) + \beta(1-\alpha)V(w, Q) + \beta\alpha Q\}. \tag{3}$$

The agent accepts the wage offer w if

$$u(b) + \beta Q < u(w) + \beta(1-\alpha)V(w, Q) + \beta\alpha Q \tag{4}$$

and otherwise rejects the offer.[3] Now observe that if (4) holds then

$$V(w, Q) = u(w) + \beta(1-\alpha)V(w, Q) + \beta\alpha Q \tag{5}$$

which implies

$$V(w, Q) = \phi u(w) + \beta\alpha\phi Q, \tag{6}$$

where $\phi = (1 - \beta(1-\alpha))^{-1}$, and we note that $0 < \alpha\phi < 1$.

We think of the optimal belief Q^* as difficult to determine, requiring as it does, a complete understanding of the wage distribution as well as the ability to compute

[3] If $u(b) + \beta Q = V(w, Q)$ the agent is indifferent between accepting the job or remaining unemployed. In this (probability zero) case, for convenience, we assume that the agent rejects the job.

fixed points. In contrast, given Q, the determination of $V(w,Q)$ is relatively straightforward: if (4) holds then $V(w,Q)$ is given by (6). The intuition for this equation can be given by rearranging (5) as

$$V(w,Q) = u(w) + \beta\left(V(w,Q) + \alpha\left(Q - V(w,Q)\right)\right).$$

This says that if accepting a job at w is optimal then its value its equal to $u(w)$ plus the discounted expected value in the coming period, which is again $V(w,Q)$ if employment continues, but must be adjusted for the "capital loss" $Q - V(w,Q)$ in value that arises if the agent becomes unemployed, which occurs with probability α.

If instead (4) does not hold, the wage offer is rejected and the agent's present value of utility is simply $u(b) + \beta Q$. We conclude that

$$V(w,Q) = \max\left\{u(b) + \beta Q, \phi u(w) + \beta\alpha\phi Q\right\}. \tag{7}$$

Thus, given perceived Q, decision-making is straightforward based on (7). We now obtain results that characterize the properties of boundedly optimal decision-making based on Q, and in the next Section we relate these results to fully optimal decision-making.

Our first result establishes the existence of a "reservation wage" $\bar{w}$ that depend on beliefs Q. Because this dependency is piece-wise it is useful to define

$$Q_\star = \frac{\phi u(w_\star) - u(b)}{\beta(1-\alpha\phi)}, \text{ where } \star \in \{\min, \max\}.$$

Proposition 1. *There is a continuous, non-decreasing function $\bar{w} : \mathbb{R} \to [w_{\min}, w_{\max}]$, which is differentiable on $(Q_{\min}, Q_{\max})$, such that $\bar{w}(Q_\star) = w_\star$ for $\star \in \{\min, \max\}$, and such that*

$$V(w,Q) = \begin{cases} u(b) + \beta Q & \text{if } Q > Q_{\max} \text{ or if } Q \in [Q_{\min}, Q_{\max}] \text{ and } w \le \bar{w}(Q) \\ \phi u(w) + \beta\alpha\phi Q & \text{if } Q < Q_{\min} \text{ or if } Q \in [Q_{\min}, Q_{\max}] \text{ and } w > \bar{w}(Q) \end{cases}. \tag{8}$$

The proof of this and all results in this Section are in Appendix A. An immediate Corollary to this proposition characterizes boundedly optimal behavior.

Corollary 1. ***(Boundedly Optimal Behavior)*** *Given beliefs Q, there exists $\bar{w}(Q) \ge w_{\min}$ such that the policy $a_t = 1$ if and only if $w_t > \bar{w}$ solves the boundedly optimal agent's problem (3).*

The optimal behavior of a boundedly rational agent with beliefs Q is characterized by a reservation wage $\bar{w}$.

Noting from proposition 1 that $\bar{w}$ depends on Q and b, we conclude this section with simple comparative statics results with respect to these variables that will be useful in Section 4. Provided that $w_{\min} < \bar{w}(Q,b) < w_{\max}$, $\bar{w}$ is implicitly defined by

$$\phi u\left(\bar{w}(Q,b)\right) + \beta\alpha\phi Q = u(b) + \beta Q. \tag{9}$$

From Assumption A we have that u is C^1 and thus

$$\frac{\partial \bar{w}}{\partial Q} = \frac{\beta(1-\alpha\phi)}{\phi u'\left(\bar{w}(Q,b)\right)} \text{ and } \frac{\partial \bar{w}}{\partial b} = \frac{u'\left(b\right)}{\phi u'\left(\bar{w}(Q,b)\right)} \tag{10}$$

which are both positive provided $u' > 0$.

Below we drop the explicit dependence of $\bar{w}$ on b except when considering cases in which b is changed.

3.2 Optimal beliefs

We now establish a link between optimal decision-making and decisions under subjective beliefs. To this end we define a map $T : \mathbb{R} \to \mathbb{R}$ by

$$T(Q) = E\left(V(\hat{w}, Q)\right) = \int_{w_{\min}}^{w_{\max}} V(\hat{w}, Q) dF(\hat{w}). \tag{11}$$

We interpret $T(Q)$ as the expected value today, induced by beliefs Q and the behavioral primitive, of having been unemployed yesterday. Lemma A.2 in Appendix A establishes that T is continuous, and is differentiable except at finite number of points, with a positive derivative strictly less than one.[4] As one would expect there is a tight link between the fixed point of this T map and optimal decision making by the agent.

Theorem 1. ***(Optimal Behavior)*** *The expected discounted utility under optimal decision-making of receiving a random wage draw, $Q^* = E\left(V^*\left(\hat{w}\right)\right)$, is the unique fixed point of the T-map (11). The policy $a = 1$ if and only if $w > \bar{w}(Q^*) \equiv w^*$ solves the optimal agent's problem (1).*

This is the standard "reservation wage" result of the McCall search model. However, theorem 1 comes with the additional interpretation that there exists a belief Q^* about the value of being unemployed such that a boundedly rational agent with beliefs Q^* behaves optimally. The explicit connection between Q^* and the agent's problem (1) arises from the observation $V^*(w) = V(w, Q^*)$, which is established in the proof of Theorem 1. This observation may then be coupled with Corollary 1, together with the equivalence of problems (2) and (3) when $Q = Q^*$.

Finally, it is convenient to adopt assumptions that result in non-trivial optimal decision-making, i.e. in which some wage offers are rejected and other wage offers are accepted: $w_{\min} < w^* < w_{\max}$. The following Proposition characterizes the parameter restrictions consistent with this assumption.

[4]Those familiar with the adaptive learning literature may be inclined to identify the condition $DT < 1$ with the E-stability condition that generally governs local stability under least-squares learning. However in a typical least-squares learning set-up, e.g. the cobweb model studied in Bray and Savin (1986) and Ch. 2 of Evans and Honkapohja (2001), E-stability concerns the parameters of the perceived law of motion of a variable viewed as exogenous to the individual agent.

Proposition 2. *If*

$$\phi\left(u(w_{\min}) - \beta(1-\alpha)\int_{w_{min}}^{w_{max}} u(\hat{w})dF(\hat{w})\right) < u(b) < \phi(1-\beta)(1-\alpha)u(w_{\max}) \quad (12)$$

then $Q_{\min} < Q^* < Q_{\max}$, *i.e.* $w_{\min} < w^* < w_{\max}$.

We omit the straightforward proof. We remark that when condition (12) holds, the comparative statics result (10) applies to Q^*. In the sequel we assume the following:

Assumption B: $u, b, w_{\min}, w_{\max}, \alpha, \beta$ and F are such that Condition (12) holds.

3.3 Learning When to Say No

We now return to considerations involving boundedly rational agents. Recall that Corollary 1 presents a reservation-wage decision rule that is optimal for given beliefs Q. For agents to learn over time in order to improve their decision-making behavior, it is necessary to update their beliefs as new data become available.[5] We adopt the "anticipated utility" perspective introduced by Kreps (1998), and frequently employed in the adaptive learning literature, in which agents make decisions based on their current beliefs Q, while ignoring the fact that these beliefs will evolve over time.[6]

As just discussed, agents update their beliefs over time as new data become available; however, we observe that if a given agent learned only from their own experience then they would update their beliefs only when they were unemployed. Because this is an implausibly extreme assumption, we introduce a social component to the adaptive learning process: we assume that in each period each agent observes a sample of wage offers received by unemployed workers and uses this sample to revise the perceived value from being unemployed. We denote by $\hat{w}_t^N = \{\hat{w}_t(k)\}_{k=1}^N$ the random sample of N wage realizations. For simplicity we assume that unemployed and employed agents use the same sample size.

Let Q_t be the value, perceived at the start of period t, of being unemployed. Note that Q_t measures the agent's perception of the value of receiving a random wage draw.[7] To update this perception the agent computes the sample mean of $V(\cdot, Q_t)$ based on his sample of wage draws. Since Q_t encodes the information from all previous wage draws, the agent updates his estimate of Q using a weighted average of Q_t with this sample mean. Formally let

$$\hat{T}(\hat{w}_t^N, Q_t) = N^{-1}\sum\nolimits_{k=1}^{N} V\left(\hat{w}_t(k), Q_t\right) \quad (13)$$

[5] Corollary 1 provides optimal decision-making given beliefs under the assumption that the separation rate is known. It would be straightforward instead to require our agents to estimate this separation rate and our asymptotic results would be unchanged.

[6] See, for example Sargent (1999), Preston (2005), Cogley and Sargent (2008).

[7] To be entirely precise, βQ_t is used as the agent's perception of the value in period t of being unemployed and therefore receiving a random wage draw in $t+1$.

denote the sample mean of $V(\cdot, Q_t)$ based on the sample $\hat{w}_t^N$. The agent is then assumed to update his beliefs at the end of period t according to the algorithm

$$Q_{t+1} = Q_t + \gamma_{t+1}\left(\hat{T}(\hat{w}_t^N, Q_t) - Q_t\right), \tag{14}$$

where $0 < \gamma_{t+1} < 1$ is specified below. Thus the revised estimate of the value of being unemployed Q_{t+1}, which is carried by the agent into the next period, adjusts the previous estimate Q_t to reflect information obtained during period t.[8]

The term $\gamma_t > 0$, known as the gain sequence, is a deterministic sequence that measures the rate at which new information is incorporated into beliefs. Two cases are of particular interest. *Constant-gain* learning sets $\gamma_t = \gamma < 1$, which implies that agents discount past data geometrically at rate $1-\gamma$. This is often used when there is the possibility of structural change, and is explored in Section 4.4 below. Under *decreasing-gain* learning $\gamma_t \to 0$ at a rate typically assumed to be consistent with assumption C below. Decreasing gain is often assumed in a stationary environment, and here provides for the possibility of convergence over time to optimal beliefs. The following assumption is made when decreasing gain is employed.

Assumption C: The gain sequence $\gamma_t > 0$ satisfies

$$\sum_{t\geq 0}\gamma_t = \infty \text{ and } \sum_{t\geq 0}\gamma_t^2 < \infty.$$

A natural example is $\gamma_t = t^{-1}$ in which data over time receives equal weight.

The following theorem is the main result of our paper.

Theorem 2. *For any Q_0, under Assumptions A, B and C, $Q_t \to Q^*$ almost surely.*

Theorem 2 establishes that in a stationary environment boundedly optimal agents will learn over time to become to make fully optimal decisions. In the next section we explore the implications of learning when there are structural changes.

A particular limiting case can help highlight the properties of this algorithm further. Consider the algorithm as $N \to \infty$. In this case, we can consider the agent as having full knowledge of the wage offer distribution. In fact, in this case we have $\hat{T}(\hat{w}_t^N, Q_t) = T(Q_t)$, where T is the T-map defined in the previous section. The evolution of beliefs is then given by

$$Q_{t+1} = Q_t + \gamma_t(T(Q_t) - Q_t).$$

Even though the agent has full knowledge of the offer distribution, she still needs to learn how to behave optimally and therefore updates beliefs in a deterministic

[8]The algorithm (14), which is standard in the adaptive learning literature, can be viewed as a special case of least-squares learning. See, for example, Ljung (1977), Marcet and Sargent (1989) and Evans and Honkapohja (2001).

manner. In fact, if the algorithm were constant gain with $\gamma_t = 1$ this would be equivalent to iterating on the agent's Bellman equation.

We view the algorithm (14) as providing a flexible model of bounded rationality that addresses several features that make dynamic optimization challenging. If N is large the sample can be viewed as revealing all needed information about the wage distribution; however, computing optimal beliefs Q^* still requires a great deal of sophistication, as noted above. Algorithm (14) provides a natural recursive updating method that can be applied to real-time decision-making. In general we assume that agents do not know the wage distribution, but even when N is small (and even in the case $N = 1$), Theorem 2 demonstrates that agents will learn Q^* over time. Finally, if structural change is a possibility then algorithm (14) with an appropriate choice of constant gain provides a way of tracking the time variation in optimal Q. We turn to these considerations now.

4 Structural Change and Transition Dynamics

Our model provides a platform for a number of comparative statics and comparative dynamics experiments; in fact, if we imagine our model populated by many agents, the comparative statics and dynamics of interesting aggregates like the unemployment rate can be examined. We first study the comparative statics and dynamics analytically, to the extent possible, under rational expectations and optimal decision-making. We decompose the changes into two terms: the direct effects hold beliefs fixed while the indirect effect comes through changes in Q^*. These two terms inform the numerical simulations allowing us to contrast the comparative dynamics under optimality and boundedly optimality.

4.1 Preliminaries

We begin by defining the variables of interest. Unemployment and duration, which will be carefully defined below, depend inversely on what we call the "hazard" rate h of leaving unemployment, i.e. the probability per period of an unemployed agent becoming employed. Given beliefs Q, as well as benefits level b and a parameter μ that will be introduced below and that will parametrize central tendency of the wage distribution, the hazard rate is

$$h = h\left(Q, b, \mu\right) = (1-\alpha)(1 - F(\bar{w}\left(Q, b\right), \mu)).$$

For a given Q, the perceived duration δ is defined to be the expected number of periods of consecutive unemployment conditional on being newly unemployed. In Appendix B it is shown that

$$\delta = \delta\left(Q, b, \mu\right) = \frac{1}{(1-\alpha)(1 - F(\bar{w}\left(Q, b\right), \mu))} = \frac{1}{h\left(Q, b, \mu\right)}. \tag{15}$$

Finally, we define u to be the unconditional probability of being unemployed; thus it must satisfy

$$u = F(\bar{w}(Q, b), \mu)u + \alpha(1 - F(\bar{w}(Q, b), \mu))u + \alpha(1 - u),$$

where the final term on the right-hand-side uses that $1 - u$ is the unconditional probability of being employed. This gives the formula

$$u = u(Q, b, \mu) = \frac{\alpha}{1 - (1 - \alpha)F(\bar{w}(Q, b), \mu)} = \frac{\alpha}{h(Q, b, \mu) + \alpha}. \tag{16}$$

The variables δ and u as just defined capture individual-level behavior. However, for a continuum of agents with fixed beliefs, we can consider how the realized aggregate unemployment rate, i.e. the cross-sectional proportion of agents who are unemployed, evolves over time from an arbitrary initial distribution of states. We observe that this distribution can be summarized by the proportion of agents with wages below the perceived cut-off $\bar{w}(Q, b)$, and further, that if this proportion is equal to $u(Q, b, \mu) \cdot F(\bar{w}(Q, b), \mu)$ then the aggregate unemployment rate is constant over time and given by $u(Q, b, \mu)$. It is also straightforward to show that, starting from any initial distribution of states, that the cross-sectional unemployment rate will converge over time to its steady-state value $u(Q, b, \mu)$.

4.2 Comparative statics under optimality

We now assume our McCall model is populated by a continuum of rational agents, and consider comparative statics associated with steady-state behavior. To compute our comparative statics, we continue to adopt Assumption B so that an interior solution exists; it follows from equation (10) that $\frac{\partial \bar{w}}{\partial Q}$ and $\frac{\partial \bar{w}}{\partial b}$ are positive.

The rational counterparts to the above definitions of h, δ and u are obtained via the observation that $Q^* = Q^*(b, \mu)$, whence

$$\begin{aligned} h^* &= h^*(b, \mu) = h(Q^*(b, \mu), b, \mu) \\ \delta^* &= \delta^*(b, \mu) = \delta(Q^*(b, \mu), b, \mu) \\ u^* &= u^*(b, \mu) = u(Q^*(b, \mu), b, \mu). \end{aligned}$$

In what follows we will compute many derivatives with respect to b and μ. When differentiating any variable other than $Q^* = Q^*(b, \mu)$, the symbol "∂" will indicate that beliefs Q are taken as fixed and the symbol "d" will indicate that beliefs Q will vary in accordance with optimality, i.e. $Q^* = Q^*(b, \mu)$. We require the following preliminary comparative statics results, which are proven in the Appendix.

Lemma 1. *If Q^* is an interior solution then $\frac{\partial Q^*}{\partial b} > 0$ and $\frac{\partial Q^*}{\partial \mu} > 0$. Hence $\frac{dw^*}{db} > 0$ and $\frac{dw^*}{d\mu} > 0$.*

The following Proposition decomposes the comparative statics of the hazard rate with respective to the benefits level into the direct and indirect effects mentioned earlier.

Proposition 3. *If Q^* is an interior solution then*

$$\frac{dh^*}{db} \equiv \frac{\partial h}{\partial b} + \frac{\partial h}{\partial Q}\frac{\partial Q^*}{\partial b} < 0,$$

with both $\frac{\partial h}{\partial b} < 0$ and $\frac{\partial h}{\partial Q}\frac{\partial Q^*}{\partial b} < 0$.

The inverse relationship between the hazard rate and both the unemployment rate and duration yields the following corollary providing comparative statics for a change in benefits.

Corollary 2. *If Q^* is an interior solution then $\frac{du^*}{db} > 0$ and $\frac{d\delta^*}{db} > 0$.*

Proposition 3 tells us that the hazard rate of leaving unemployment is decreasing in unemployment benefits. This effect is decomposed into direct effect and indirect effects. $\frac{\partial h}{\partial b}$ captures the direct effect: even if agents do not update their beliefs they will still react to an increase in benefits by raising their reservation wage. Proposition 3 tells us that a rational agent would respond even further by taking into account that higher unemployment benefits also raise the value of Q^*. This is the indirect effect. While the hazard rate for the rational agents exhibits no dynamics, i.e. jumps from the old steady-state value to the new one, under learning the hazard rate evolves over time as beliefs Q are updated. For this reason, the indirect effects are not initially incorporated into the boundedly rational agents' hazard rate.

Turning now to the impact of a change in the median wage (or some other measure of central tendency), some assumptions are needed on the distribution of the wage draws. Thus let $\mathcal{I}$ be a connected subset of the reals and let $\{F(\cdot,\mu)\}_{\mu\in\mathcal{I}}$ be a family of distributions that is C^1 in the index parameter μ. Here, μ is intended to capture some abstract measure of central tendency. We assume that μ orders this family of distributions by (first-order) stochastic dominance:

$$\mu_1 \leq \mu_2 \implies F(w,\mu_1) \geq F(w,\mu_2).$$

We have the following:

Proposition 4. *Assume Q^* is an interior solution. Then*

$$\frac{dh^*}{d\mu} \equiv \frac{\partial h}{\partial \mu} + \frac{\partial h}{\partial Q}\frac{\partial Q^*}{\partial \mu} > 0 \iff \frac{\partial \bar{w}}{\partial Q}\frac{\partial Q^*}{\partial \mu} < -\frac{\partial F/\partial \mu}{dF}. \tag{17}$$

The condition in (17) captures competing effects of an increase in the median wage, and arises from the following computation:

$$\frac{\partial h}{\partial \mu} + \frac{\partial h}{\partial Q}\frac{\partial Q^*}{\partial \mu} = -(1-\alpha)\left(\frac{\partial F}{\partial \mu} + \frac{\partial \bar{w}}{\partial Q}\frac{\partial Q^*}{\partial \mu}dF\right).$$

Intuitively, the rise in μ increases the proportion of wage offers that are higher than w^* (direct effect), while at the same time raising w^* and hence the proportion of agents who reject wage offers (indirect effect). When the condition in (17) holds, the direct effect dominates; however, the direct and indirect effects of the change in μ are competing, so the hazard rate of boundedly rational agents initially overshoots the new steady state value.

Specific families of distributions with specific dependences on the index parameter μ provide additional insight and are useful for the simulations presented below. We consider two families of distributions. The first is simply a collection of translations of our given distribution F. Thus $\mathcal{I}$ is a small interval about zero and

$$F(\bar{w},\mu) = \int_{w_{\min+\mu}}^{\bar{w}} dF(w-\mu). \tag{18}$$

In this case, an increase in μ may be interpreted as an increase in the median wage, and it is immediate that μ orders the family of distributions by stochastic dominance.[9]

The second collection of distributions is the parametric family of lognormal distributions with shape parameters[10] μ and s, which implies a median wage e^{μ}. In this case, $\mathcal{I} = \mathbb{R}$ and distributions are given by

$$F(\bar{w},\mu) = \frac{1}{s\sqrt{2\pi}} \int_0^{\bar{w}} \frac{1}{w} \cdot e^{-\frac{(\mu-\log(w))^2}{2s^2}} dw. \tag{19}$$

That for fixed s, the shape parameter μ orders the lognormal family of distributions by stochastic dominance is established in Levy (1973). Again, the inverse relationship between the hazard rate and both the unemployment rate and duration yields the following corollary to Proposition 3 providing comparative statics for a change in μ.

Corollary 3. *Assume Q^* is an interior solution.*

1. *If $F(\bar{w},\mu)$ is given by (18) then $\frac{du^*}{d\mu} < 0$ and $\frac{d\delta^*}{d\mu} < 0$ if and only if the level change of the reservation wage resulting from a change in the median wage is less one*

2. *If $F(\bar{w},\mu)$ is given by (19) then $\frac{du^*}{d\mu} < 0$ and $\frac{d\delta^*}{d\mu} < 0$ if and only if the elasticity of the reservation wage with respect to the median wage is less than one.*

4.3 Comparative dynamics under optimality

With rational agents, only the unemployment rate experiences non-trivial transition dynamics; the hazard rate and duration for the newly unemployed simply jump to

[9] Here we are abusing notation somewhat: If F has a single argument then it refers to the given distribution and if F has two arguments then it references to the shifted distribution. Thus $dF(w,\mu) = dF(w-\mu)$.

[10] The random variable x is lognormally distributed with shape parameters μ, s provided $\log(x) \sim N(\mu, s^2)$.

their new steady state levels. The same would be true for boundedly optimal agents if their beliefs Q were constant over time; however, under learning the evolution over time of beliefs induces transition dynamics in the hazard rate. To examine unemployment dynamics it is helpful to define the notion of a "quit." We say that an agent employed in time $t-1$ quits in time t, and thereby becomes unemployed, if his wage in time $t-1$ is less that w_t^*. Here the t subscript allows for variations in the optimal reservation wage induced by structural change. We observe that quits can only occur in case a structural change between periods $t-1$ and t results in $w_t^* > w_{t-1}^*$. Therefore, to simplify our analysis we will assume that a structure change at time 0 occurs only after a long period of stability so that the economy has reached a long run steady state. We focus on the dynamics of rational agents but, as in the previous section, we decompose changes in unemployment into direct and indirect effects to shed light on the unemployment dynamics with boundedly rational agents.

Under this assumption, let w_{-1} denote the wage of individual drawn randomly in period -1 from the pool of employed individuals. The probability that this individual quits in period 0 is given by

$$q_0 = q\left(w_0^*\right) = \frac{\max\left\{0, F(w_0^*, \mu_{-1}) - F(w_{-1}^*, \mu_{-1})\right\}}{1 - F(w_{-1}^*, \mu_{-1})},$$

where we have exploited that the long run distribution of wages will be the distribution of wage offers, $F(\cdot, \mu_{-1})$ truncated at the reservation wage w_{-1}^*. The time subscript on μ is present to indicate that even if μ captures the structural change, the distribution functions are evaluated at the "old" value of μ. This reflects that the old value of μ characterizes the distribution from which w_{-1} is randomly drawn. Interpreted cross-sectionally, q_0 is the proportion of agents employed in time -1 who quit in time 0. Now let u_t be the proportion of agents who are unemployed in period t . Noting that $1-u$ is the proportion of employed agents, and that $1-h$ is the probability that an unemployed agent remains unemployed, the dynamics of u_t may be written

$$u_t = (1-h_t)\, u_{t-1} + (\alpha + (1-\alpha)q_t)(1-u_{t-1}). \tag{20}$$

In case of no structural change for all t then $q_t = 0$ and the unemployment rate u_t converges to the steady-state unemployment level

$$u^* = \frac{\alpha}{h+\alpha} = \frac{\alpha(1-u^*)}{h},$$

where the second equality will facilitate matters below.

For comparison with the impact effect of structural change on the unemployment rate, it is useful to recognize the decomposition

$$\frac{du^*}{d\star} \equiv \frac{\partial u^*}{\partial \star} + \frac{\partial u^*}{\partial Q}\frac{\partial Q^*}{\partial \star}, \tag{21}$$

for a change in the structural parameter $\star \in \{b, \mu\}$. We can now evaluate the change in unemployment driven by a change in structural parameters $\star$ assuming that the economy is initially in steady state. Differentiation of (20) at $t = 0$ yields

$$\frac{du_0}{d\star} = -u^* \frac{dh^*}{d\star} + (1-\alpha)(1-u^*) \frac{dq}{dw^*} \cdot \frac{dw^*}{d\star}. \tag{22}$$

It is important to emphasize here that we are differentiating q at the previous steady state reservation wage, and while q is not differentiable at this point it is Gateaux differentiable with

$$dq = \begin{cases} \frac{dF(w^*_{-1}, \mu_{-1})}{1-F(w^*_{-1}, \mu_{-1})} dw^* & \text{if } dw^* \geq 0 \\ 0 & \text{if } dw^* < 0 \end{cases}$$

$$\frac{dq}{dw^*} \cdot \frac{dw^*}{d\star} = \begin{cases} \frac{dF(w^*_{-1}, \mu_{-1})}{1-F(w^*_{-1}, \mu_{-1})} \frac{dw^*}{d\star} & \text{if } \frac{dw^*}{d\star} \geq 0 \\ 0 & \text{if } \frac{dw^*}{d\star} < 0 \end{cases}$$

Thus, noting that by Lemma 1 $\operatorname{sign}(dw^*) = \operatorname{sign}(d\star)$, we have the following:

Proposition 5. *The differential of unemployment with respect to a change in structural parameter $\star$ is given by*

$$du_0 = \begin{cases} \left(\frac{1}{u^*} \frac{du^*}{d\star} + \left(\frac{1-\alpha}{\alpha}\right) u^* \cdot h_\star\right) d\star & \text{if } d\star \geq 0 \\ \frac{\alpha}{u^*} \frac{du^*}{d\star} d\star & \text{if } d\star < 0. \end{cases}$$

Applying Proposition 5 to our two examples of structural change allows us to highlight the asymmetry in the response of unemployment and the role of beliefs. Let us begin with a change in unemployment benefits b. $\frac{\partial F^*}{\partial b} = 0$ implies that

$$du_0 = \begin{cases} \frac{1}{u^*} \frac{du^*}{db} db & \text{if } db \geq 0 \\ \frac{\alpha}{u^*} \frac{du^*}{db} db & \text{if } db < 0. \end{cases}$$

As $\frac{1}{u^*}$ is much larger than one we can conclude that an unexpected increase in unemployment benefits will result in an initial spike in unemployment many times of that of the increase in steady state unemployment. On the other hand α/u^* is necessarily less than one, which implies that a decrease in benefits will result in a fall in unemployment smaller than the fall in steady state. In both cases, the initial change in unemployment can be decomposed into indirect and direct effects by decomposing $\frac{du^*}{db}$ via equation (21). Proposition 3 allows us to conclude both the direct and indirect effects move in the same direction and thus we would expect the response of the boundedly rational agents to be smaller.

The predictions of Proposition 5 for a change in μ are ambiguous. We assume for this discussion that the reservation wage assumptions of Corollary 3 hold, so that $du^*/d\mu < 0$. If μ decreases then w^* decreases, whence there is no change in quits q. It follows that u_0 unambiguously rises in accordance with the second line of Proposition 5. On the other hand, an increase in μ will result in an increase in w^*. Thus, while wage draws will improve, putting downward pressure on unemployment, some workers will quit due to the rise in the reservation wage. This tension renders the total impact on u_0 ambiguous in the fully optimal case. In contrast, if we focus on the direct effect then equation (22) implies [11]

$$\frac{\partial u_0}{\partial \mu} = -u^* \frac{\partial h}{\partial \mu}.$$

Since $\frac{\partial h}{\partial \mu} > 0$, it follows that $\frac{\partial u_0}{\partial \mu} < 0$ unambiguously, and this does not depend on whether the reservation wage assumptions of Corollary 3 hold. We would, therefore, expect unemployment to decrease in response to an increase in the reservation wage if agents were boundedly rational.

Finally, following the structural change, $q_t = 0$ for $t \geq 1$ and the unemployment dynamics under rationality may be written

$$u_t - u^* = (1 - \alpha - h^*)\,(u_{t-1} - u^*)\,, \tag{23}$$

where h^* and u^* correspond to their new steady-state values. Since the hazard rate is bounded above by $1 - \alpha$, equation (23) implies geometric monotonic convergence of unemployment to its new steady-state level.

4.4 Comparative dynamics under bounded optimality

We now use numerical methods to study comparative dynamics in our model. The simulations in this Section are based on the following specification. All simulations are conducted with a constant gain of $\gamma = 0.1$. Utility is $CRRA$ with risk aversion parameter $\sigma > 0$, and the exogenous wage distribution is taken to be lognormal[12] with parameters $\mu, s > 0$, yielding a median wage of e^{μ} and variance $e^{2\mu+s^2}\left(e^{s^2} - 1\right)$. In our calibration we set $\mu = 11.0$, $s = 0.25$ and $\sigma = 4.5$. In addition we set $\beta = 0.99$ and the separation rate is set at $\alpha = 0.025$.[13] Our value for σ is higher than typically used in macroeconomic models, but consistent with the range considered in asset-pricing models. The baseline value of μ corresponds to a median household wage of approximately $60,000$, close to the US value in dollars in 2016. For the choice of s, what is relevant for our model is the distribution of wage income faced by the individual agent, i.e. not a measure of the population wage distribution.[14] At our

[11] For changes in μ, changes in the reseervation wage only come through changes in beliefs so $\frac{\partial w^*}{\partial \mu} = 0$.

[12] Although lognormal does not have impose $w_{\min} > 0$ or $w_{\max} < \infty$, this is numerically indistinguishable from setting $w_{\min}$ small and $w_{\max}$ large.

[13] The simulation results are robust to higher values of β, thus allowing for various interpretations of the time period.

[14] Our value for s is broadly consistent with the literature. For example, p. 576 of Greene (2012) using a pooled LS estimate of a log wage equation controlling for a number of individual specific

baseline value $s = 0.25$ the interquartile income range is $50,583$ to $70,871$. The lowest decile ends at $w = 43,460$ and the highest decile begins at $w = 82,486$. Finally, the baseline unemployment benefit parameter is set at 50% of the median wage. We interpret our calibration as capturing the experience of an individual interacting in a local labor market populated by individuals with similar characteristics.

We begin with a change in the benefit b. In all of our simulations we track 6000 agents, i.e. the size of this local local labor market, who experience an unexpected structural change at time $t = 50$; and in all Figures, the horizontal (red) dotted lines represent the pre-shock optimal values and the horizontal (blue) dashed lines represent post-shock optimal values.[15] Figure 1 presents a simulation with a 20% increase in benefits. For fixed beliefs Q, an increase in benefits db results in an increase in the instantaneous return $u'(b)\,db$ to being unemployed, thereby raising the reservation wage. This the direct effect emphasized in the previous Section. The corresponding indirect effect of a rise in benefits is that it also raises the optimal present value Q^* of being unemployed. For the rational agents both effects are instantaneous, whereas for the boundedly optimal agents, the initial impact on the reservation wage is only through the increase in the instantaneous return, with the impact from changes in Q developing over time.

The decomposition into direct and indirect effects is evidenced in the lower left-hand panel of Figure 1: each of the paths provides the realized time series of the wage cut-off for a given agent. For the first 50 periods these paths are distributed around the pre-shock optimal cut-off wage – the distribution reflects the evolving beliefs of different agents as determined by their idiosyncratic sample draws. At time $t = 50$ all paths exhibit a sharp increase in the wage cut-off due to the rise in b, which uniformly affects all agents. Subsequently over time, as evidenced in the upper right-hand panel, agents' beliefs converge to a distribution around the new optimal value of Q, and the distribution of corresponding wage cutoffs evolves to a distribution around the new optimal wage cut-off.[16]

Turning to unemployment duration, the time series presented in the upper-left panel gives, at each point in time, the realized cross-agent average, conditional on being newly unemployed, of the number of periods until the agent is next employed.

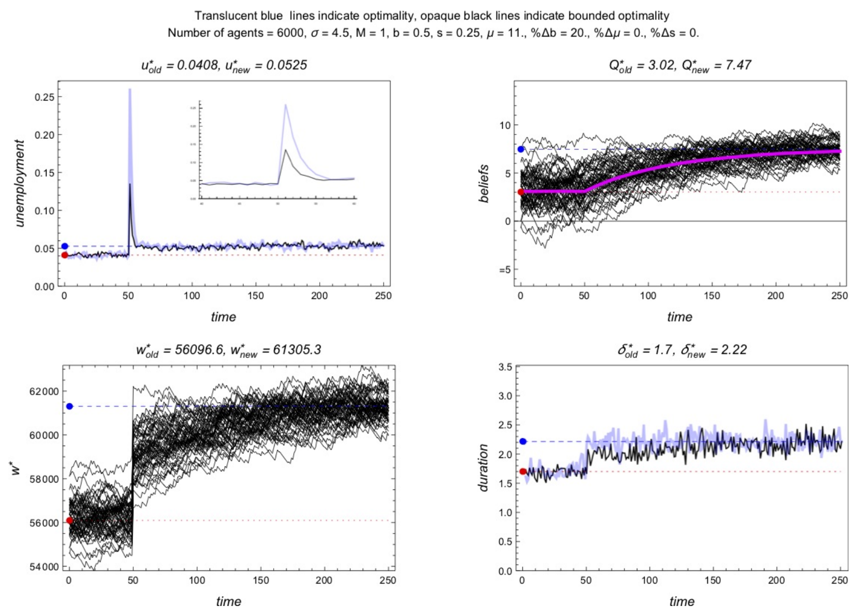
---

characteristics, obtains a residual variance of 0.146, i.e. $s = 0.382$. Krueger et al. (2016), estimate a log-labor earnings process with persistent and transitory shock. They find that the variance of the transitory shocks, which are the shocks more relevant for our model, is 0.0522, i.e. $s = 0.23$. The qualitative features of the simulations are robust to values of s across this range.

[15] All simulations are initialized by providing boundedly rational agents with beliefs in a small neighborhood of the optimal value of Q, and with the percent of agents identified as unemployed corresponding to the rational model's steady-state unemployment rate. To eliminate transient dynamics the model is run for a large number of periods before our simulation begins.

[16] It is interesting to note that the cross-sectional variation of beliefs decreases after the policy change. This behavior reflects that a rise in benefits leads to an increase in the unemployment rate, which increases the proportion of value function realizations determined by the nonstochastic component of the maximization problem: see equations (7) and (8).

Figure 1: An increase in benefits



For rational agents the expected unemployment duration for the newly unemployed jumps to the new steady-state duration level, whereas, because their $\bar{w}$ does not fully adjust immediately, boundedly rational agents are initially more likely to take jobs, leading to a more gradual adjustment of the duration.

Finally, we consider the unemployment time series, in the lower right panel, which dramatically illustrates the discrepancy in behavior of the optimal and boundedly optimal agents at the time of the policy change. As noted in the previous Section, an increase in benefits leads to an increase in the rational-agent steady-state unemployment rate. The translucent (blue) path identifies the unemployment rate associated with the rational-agents simulation.[17] This time series exhibits a very large spike at the time of the shock, a quintupling in fact, which reflects the impact effect identified in the discussion following Lemma 5. This spike can be explained by the behavior of the associated wage cut-off: because optimal agents experience both the direct and indirect effects at the instant of the change in b, their wage cut-off rises immediately to the new optimal level, an increase of over 9%, which causes a dramatic rise in unemployment resulting from previously employed agents not accepting their wage offers. The behavior of the boundedly optimal agents is similarly explained, but is muted by the failure of the indirect effect to materialize immediately. The inset of the lower right panel shows a more detailed view of the same simulation near the time of the policy change.

[17]The stochastic fluctuations in the aggregate time series arise from the fact that our population of agents is finite (6000).

Figure 2, which examines a decrease in the benefit rate, exhibits a different dynamic response of the unemployment rate: it falls, as one would expect, but there is no overshooting spike in either the rational or boundedly rational case. This is easily understood: when benefit rates rise, employed agents with low wages immediately quit their jobs to capture the increased benefit of being unemployed; however, when benefits fall, all employed agents have increased incentives to retain their jobs and unemployed agents are willing to accept lower wages, but not to the extent that overshooting is implied.

Figure 2: A decrease in benefits

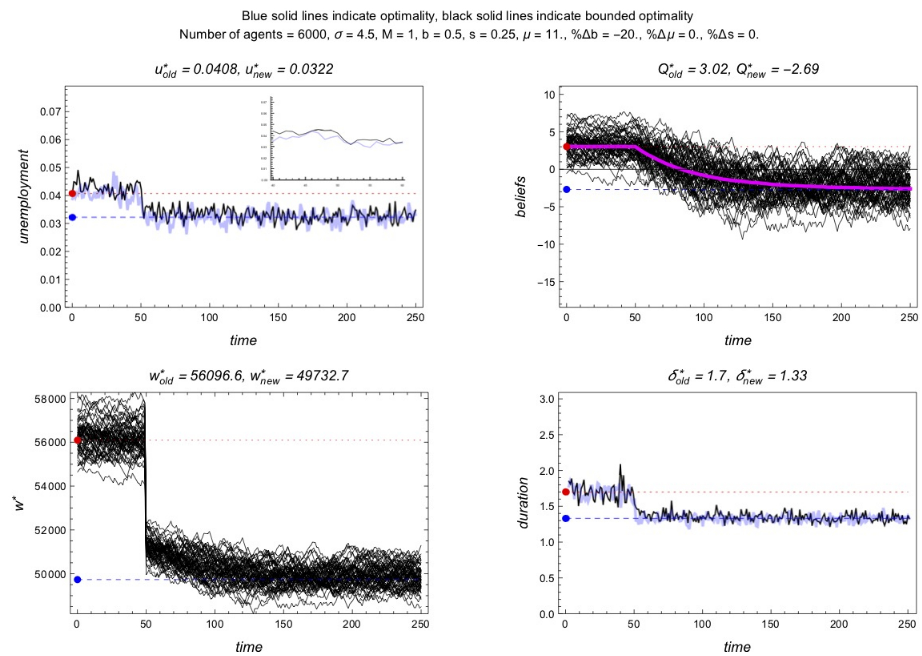


Figure 3 examines the effect of an increase in μ of 2.5%, which leads to an increase in the median wage from approximately 60,000 to 78,800. This admittedly dramatic increase conveniently induces a fall of the steady-state unemployment rate by one percentage point, given our calibration of the model. This increase in μ leads to a large increase in w^*, which is entirely due to the large increase in Q^*, and which leads to a dramatic spike in the unemployment rate in the rational-agent model. Intuitively, at the time of the shock, employed rational agents with relatively low wages quit their jobs in order to obtain new wage draws from the improved wage distribution.[18]

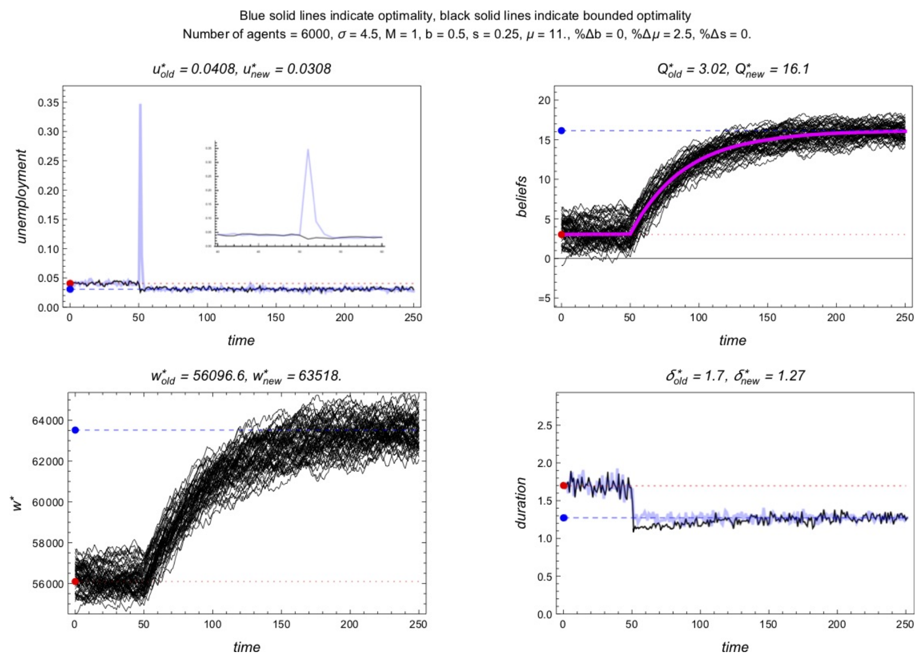
No such spike is observed in the model with boundedly rational agents, which again reflects that the reservation wage is only responding to changes in beliefs. Boundedly rational agents are not aware of the shock and only learn about it over time as new

[18]As indicated in the paragraph following Lemma 5, the positive spike to unemployment seen in Figure 3 is calibration specific; in some extreme cases the sign of the impact effect can be reversed.

wage draws are obtained. As they adjust their beliefs accordingly, their wage cutoffs increase toward the new optimal steady-state level.

The unemployment dynamics exhibit overshooting, most easily seen in the duration panel. Specifically, the duration of unemployment dips below the new steady-state level, before converging to it. This behavior is explained by the gradual response of the reservation wage: at initial impact, the reservation wage does not change and wage draws are more likely to be higher, which leads to a greater proportion of agents accepting the offers.

Figure 3: An increase in median wage



In contrast to both the rational case when wages fall, and to the boundedly optimal case when wages rise, the unemployment time series for boundedly optimal agents experiencing a surprise fall in wages exhibits considerable overshooting: see both the unemployment rate and duration panels of Figure 4. This overshooting reflects the failure of learning agent to recognize the deterioration in the labor market. Specifically, the reservation wages for boundedly optimal agents do not change on impact of the wage-distribution shock, and so these agents reject a much higher proportion of the new wage offers, thus leading to a sharp rise in both the unemployment rate and duration. As new data on the wage distribution are obtained, the agents adjust their beliefs and associated wage cutoffs, causing the unemployment rate and duration to converge to their new steady-state levels.

The asymmetric effects of the boundedly rational agents in response to changes in the median wage suggests a corresponding asymmetry in the business cycle. To

Figure 4: A decrease in median wage

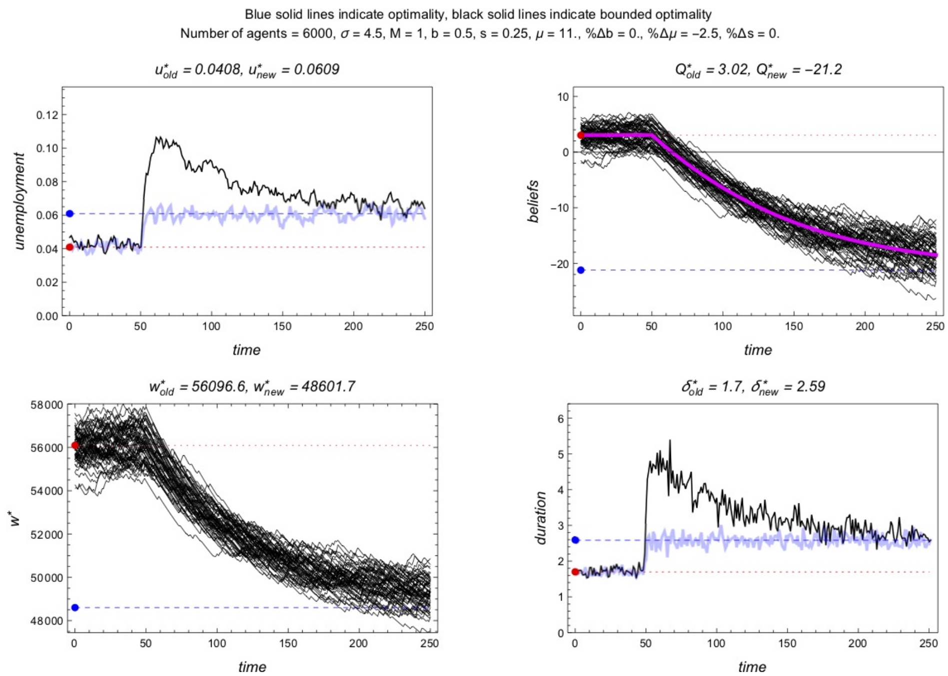
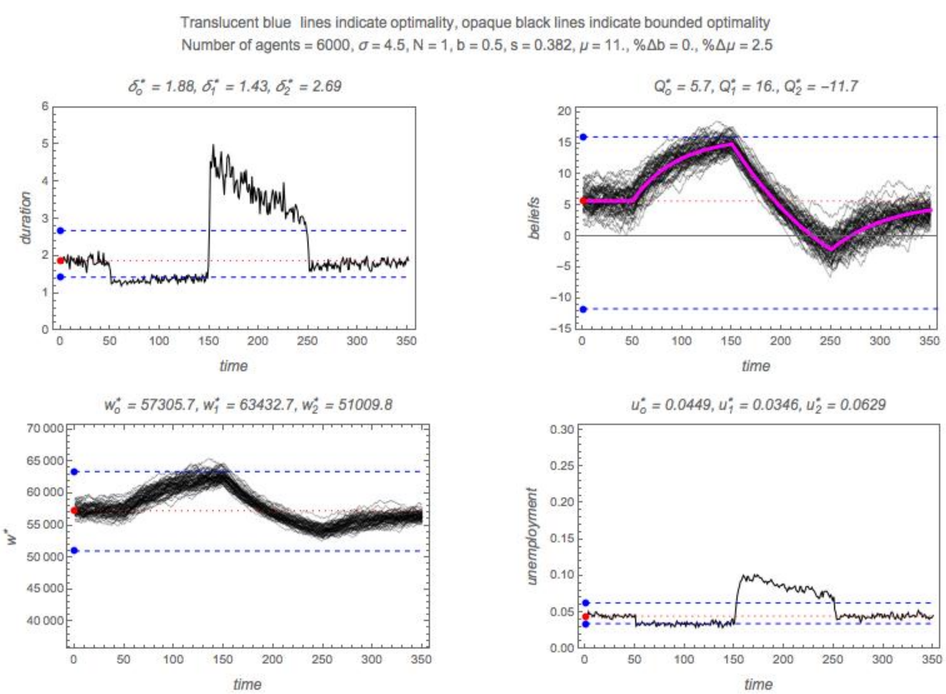


Figure 5: Impact of a wage cycle



illustrate this potential within the context of our partial equilibrium model, we consider the following experiment: at period $t = 50$ the median wage increases by 2.5%, then, in period $t = 150$, falls to 2.5% below the original steady state; finally, in period $t = 250$ the wage returns to the original steady state. Figure 5 presents the results for the model under learning. Observe, for example, the bottom right panel providing the unemployment time series: at the time of the positive shock to wages the unemployment path falls quickly to the new steady-state level, whereas the negative shock to wages results in dramatic and prolonged overshooting. This behavior could be reflected in business-cycle patterns which tend to display episodes of relatively steady moderated growth punctuated by sharp declines and slow recoveries.

5 Conclusions

We consider boundedly optimal behavior in a well known partial-equilibrium model of job search. Boundedly optimal decision-making depends on a univariate sufficient statistic that summarizes the perceived value to the job-seeker of receiving a random wage draw. Following the adaptive learning literature, agents update their perceived values over time based on their current perceptions and observed wage draws. We show that, under natural assumptions, this learning algorithm is globally stable: given any initial perception, our boundedly optimal agents learn over time to make optimal decisions.

Using numerical simulations we consider structural change and compare the dynamics of an economy populated by fully rational agents to those of an economy populated by boundedly optimal agents. More specifically we consider changes in the unemployment benefit level and changes in the wage distribution. We find that either an increase in benefits or in the median wage causes a large spike in unemployment under rational expectations, which under learning is dampened (rise in benefits) or nonexistent (rise in median wage). Further, a fall in the median wage causes significant and persistent overshooting of the unemployment rate for boundedly optimal agents above the new, higher rational level.

Appendix A: Proofs of results in Section 3

Proof of Proposition 1. First, observe that the agent rejects the wage offer w if and only if

$$\phi u(w) \leq u(b) + \beta(1-\alpha\phi)Q. \tag{24}$$

The argument is completed by addressing the following three cases:

1. If $Q > Q_{\max}$ then condition (24) always holds; thus $\bar{w}(Q) = w_{\max}$, the agent rejects any offer and receives $u(b) + \beta Q$.

2. If $Q < Q_{\min}$ then condition (24) never holds; thus $\bar{w}(Q) = w_{\min}$, the agent accepts any offer w and receives $\phi u(w) + \beta\alpha\phi Q$.

3. Finally, if $Q_{\min} \leq Q \leq Q_{\max}$ then

   $$\phi u\left(w_{\min}\right) \leq u(b) + \beta(1-\alpha\phi)Q \leq \phi u\left(w_{\max}\right). \tag{25}$$

   Since $u'(w) > 0$ it follows that for each $Q \in [Q_{\min}, Q_{\max}]$ there is a unique $\bar{w}(Q) \in [w_{\min}, w_{\max}]$ such that

   $$\phi u(\bar{w}(Q)) = u(b) + \beta(1-\alpha\phi)Q,$$

   and further that, in this case, condition (24) holds if and only if $w \leq \bar{w}(Q)$.

   It remains to show that, so defined, $\bar{w}$ is differentiable on $(Q_{\min}, Q_{\max})$. Since u is C^2, by the implicit function theorem, it follows that for each $Q \in (Q_{\min}, Q_{\max})$ there is an open set $U(Q) \subset (Q_{\min}, Q_{\max})$ and a differentiable function $g_Q : U(Q) \to [w_{\min}, w_{\max}]$ such that for all $Q' \in U(Q)$,

   $$\phi u(g_Q(Q')) = u(b) + \beta(1-\alpha\phi)Q',$$

   and, further, by uniqueness of $\bar{w}(Q')$, we may conclude that $\bar{w} = g_Q$ on $U(Q)$. Since the $U(Q)$ cover $(Q_{\min}, Q_{\max})$ the proof is complete. ■

To establish Theorem 1 we need the following technical result:

Lemma A.1. *If $f : \mathbb{R} \to \mathbb{R}$ is continuous, if f is differentiable except at perhaps a finite number of points, and if the derivative of f, when it exists, is positive except at perhaps a finite number of points, then f is strictly increasing.*

Proof: In the context of this proof, we say that x_0 is *anomalous* if either $f'(x_0)$ does not exist or $f'(x_0) \leq 0$. We begin by assuming f has only one anomalous point x_0. Because the derivative is positive for $x \neq x_0$, it suffices to show that if $x < x_0$ then $f(x) < f(x_0)$ and if $x > x_0$ then $f(x) > f(x_0)$. Suppose $x < x_0$. By the mean value theorem applied to $[x, x_0]$, which requires that f be continuous on $[x, x_0]$ and differentiable on (x, x_0), there exists $x^* \in (x, x_0)$ such that

$$\begin{aligned}
\frac{f(x_0) - f(x)}{x_0 - x} &= f'(x^*), \text{ or} \\
f(x_0) - f(x) &= f'(x^*)\,(x_0 - x) > 0.
\end{aligned}$$

An analogous argument holds if $x > x_0$. Finally, this argument is easily generalized to account for a finite number of anomalous points. $\blacksquare$

The following Lemma, which is referenced in the main text, establishes important properties of the T-map, including an upper bound on its derivative.

Lemma A.2. *The map given by (11) is continuous on* $\mathbb{R}$, *differentiable everywhere except possibly* $Q_{\min}$ *and* $Q_{\max}$, *and* $0 < DT \leq \beta < 1$ *whenever it exists.*

Proof. Using Proposition 1, direct computation yields the following formulation of the T-map:

$$T(Q) = \begin{cases} \alpha\beta\phi Q + \phi \int_{w_{\min}}^{w_{\max}} u(w) dF(w) & \text{if} \quad Q < Q_{\min} \\ \begin{array}{l} (u(b) + \beta Q) F(\bar{w}(Q)) + (1 - F(\bar{w}(Q))) \beta\alpha\phi Q \\ +\phi \int_{\bar{w}(Q)}^{w_{\max}} u(w) dF(w) \end{array} & \text{if} \quad Q_{\min} \leq Q \leq Q_{\max} \\ u(b) + \beta Q & \text{if} \quad Q > Q_{\max} \end{cases}.$$

Clearly $DT(Q) > 0$. It further follows from Proposition 1 that the map T is continuous on $\mathbb{R}$ and differentiable everywhere except possibly $Q_{\min}$ and $Q_{\max}$. Next we compute an upper bound on DT. If $Q < Q_{min}$ then $DT(Q) = \beta\alpha\phi < \beta$, where the inequality follows from $\alpha\phi \in (0, 1)$. If $Q > Q_{\max}$ then $DT(Q) = \beta$. Finally, if $Q_{\min} < Q < Q_{\max}$ we may compute

$$\begin{aligned} DT(Q) &= (u(b) + \beta Q) \, dF(\bar{w}) \frac{\partial \bar{w}}{\partial Q} + \beta F(\bar{w}) - (\phi u(\bar{w}) + \beta\alpha\phi Q) dF(\bar{w}) \frac{\partial \bar{w}}{\partial Q} \\ &\quad + (1 - F(\bar{w})) \beta\alpha\phi \\ &= \beta \left(F(\bar{w}(Q)) + (1 - F(\bar{w}(Q))) \alpha\phi \right) < \beta, \end{aligned}$$

where the second equality exploits the definition of $\bar{w}$. $\blacksquare$

Proof of Theorem 1. We begin the proof by establishing that the T-map has a unique fixed point. Let

$$\hat{Q} \leq \min \left\{ \frac{\phi u(w_{\min})}{1 - \alpha\beta\phi}, Q_{\min} \right\}.$$

We claim that $T(\hat{Q}) > \hat{Q}$. Indeed,

$$T(\hat{Q}) = \alpha\beta\phi\hat{Q} + \phi \int_{w_{\min}}^{w_{\max}} u(w) dF(w) > \alpha\beta\phi\hat{Q} + \phi u(w_{\min}) \geq \hat{Q}.$$

Next, let $h : \mathbb{R} \to \mathbb{R}$ be defined as

$$h(Q) = T(\hat{Q}) + \beta(Q - \hat{Q}).$$

We claim $Q \geq \hat{Q}$ implies $h(Q) \geq T(Q)$. Indeed let $H(Q) = h(Q) - T(Q)$. Then H is continuous and $H'(Q) > 0$ except perhaps at $Q_{\min}$ and $Q_{\max}$. Thus by Lemma A.1, H is strictly increasing. The claim follows from the fact that $H(\hat{Q}) = 0$.

Finally let $\check{Q} \equiv (1-\beta)^{-1}\left(T(\hat{Q}) - \beta\hat{Q}\right)$. Then

$$Q \geq \check{Q} \Rightarrow h(Q) < Q \Rightarrow T(Q) < Q.$$

Thus we have $T(\hat{Q}) > \hat{Q}$ and $T(\check{Q}) < \check{Q}$. Since T is continuous, the existence of a fixed point Q^* is guaranteed by the intermediate value theorem. Finally, let $S(Q) = Q - T(Q)$. Then S is continuous and $S'(Q) > 0$ except perhaps at $Q_{\min}$ and $Q_{\max}$. Thus by Lemma A.1, S is strictly increasing, from which it follows that the fixed point of T is unique.

Now we turn to connecting Q^* to the Bellman functional equation (2), which we repeat here for convenience:

$$\begin{aligned} V(w) &= \max_{a\in\{0,1\}} u\left(c(a,w)\right) + \beta E\left(V(w')|a,w\right) \\ & \quad w' = g(w,a,\hat{w},s). \end{aligned}$$

The binary nature of the choice variable makes this problem accessible. Specifically,

$$\begin{aligned} E\left(V(w')|0,w\right) &= \int V(\hat{w})dF(\hat{w}) \\ E\left(V(w')|1,w\right) &= (1-\alpha)V(w) + \alpha\int V(\hat{w})dF(\hat{w}). \end{aligned}$$

It follows that

$$a = 0 \implies V(w) = u(b) + \beta\int V(\hat{w})dF(\hat{w}) \tag{26}$$

$$a = 1 \implies V(w) = u(w) + \beta(1-\alpha)V(w) + \alpha\beta\int V(\hat{w})dF(\hat{w}), \text{ or}$$

$$a = 1 \implies V(w) = \phi u(w) + \phi\alpha\beta\int V(\hat{w})dF(\hat{w}), \tag{27}$$

where $\phi = (1-\beta(1-\alpha))^{-1}$. We conclude that the Bellman functional equation may be rewritten as

$$V(w) = \max\left\{u(b) + \beta\int V(\hat{w})dF(\hat{w}), \phi u(w) + \phi\alpha\beta\int V(\hat{w})dF(\hat{w})\right\}. \tag{28}$$

Now define $\tilde{Q} = \int V(\hat{w})dF(\hat{w})$, which may be interpreted as the value of having a random draw from the exogenous wage distribution. Then equation (28) becomes

$$V(w) = \max\left\{u(b) + \beta\tilde{Q}, \phi u(w) + \phi\alpha\beta\tilde{Q}\right\}, \tag{29}$$

from which it follows that

$$\tilde{Q} = \int V(w)dF(w) = \int\left(\max\left\{u(b) + \beta\tilde{Q}, \phi u(w) + \phi\alpha\beta\tilde{Q}\right\}\right)dF(w). \tag{30}$$

Using Proposition 1 we may write

$$\begin{aligned}
&\int \left(\max\left\{u(b)+\beta\tilde{Q}, \phi u(w)+\phi\alpha\beta\tilde{Q}\right\}\right) dF(w) \\
= \quad & (u(b)+\beta\tilde{Q})F\left(\bar{w}\left(\tilde{Q}\right)\right)+\phi\int_{\bar{w}(\tilde{Q})}^{w_{\max}} u(w)dF(w)+\phi\alpha\beta\tilde{Q}\left(1-F\left(\bar{w}\left(\tilde{Q}\right)\right)\right).
\end{aligned}$$

We conclude that equation (30) can be written

$$\tilde{Q}=(u(b)+\beta\tilde{Q})F\left(\bar{w}\left(\tilde{Q}\right)\right)+\phi\int_{\bar{w}(\tilde{Q})}^{w_{\max}} u(w)dF(w)+\phi\alpha\beta\tilde{Q}\left(1-F\left(\bar{w}\left(\tilde{Q}\right)\right)\right)=T(\tilde{Q}),$$

where the last equality follows from the definition of T. Since the T-map has a unique fixed point Q^*, we conclude that $\tilde{Q}=Q^*$. By equation (29) $\tilde{Q}$, and hence Q^*, uniquely identifies V, the solution to the Bellman system. It follows from equation (7) that $V(w)=V(w,Q^*)$. Finally, Corollary 1 implies $w^*=\bar{w}(Q^*)$. ■

To prove Theorem 2, we require the following technical Lemma:

Lemma A.3. *Suppose that γ_n is a sequence of positive numbers satisfying $\sum_n \gamma_n^2<\infty$. The following are equivalent:*

Lemma 2. *a. $\sum_n \gamma_n=\infty$.*

b. There exists $\lambda>0$ such that $\prod_n(1-\lambda\gamma_n)=0$.

c. $\prod_n(1-\lambda\gamma_n)=0$ for all $\lambda>0$.

Proof. Denote by $\left\{\gamma_n^N\right\}$ the N-tail of $\{\gamma_n\}$, that is, $\gamma_n^N=\gamma_{N+n}$. It will be helpful to observe that since $\gamma_n\to 0$, given $\varepsilon>0$ there is an $N>0$ so that $\gamma_n^N<\varepsilon$ for all $n>0$.

$(a\Rightarrow c)$. Let $\lambda>0$ and choose $N_2(\lambda)>0$ so that $\lambda\gamma_n^{N_2}<1$ for all $n>0$. By the concavity of the logarithm, we have that

$$\log\left(1-\lambda\gamma_n^{N_2}\right)<-\lambda\gamma_n^{N_2}.$$

Now define

$$P_M^{N_2}(\lambda)=\prod_{n=1}^{M}(1-\lambda\gamma_n^{N_2}),$$

and observe that

$$\log P_M^{N_2}(\lambda) \quad < \quad -\lambda\sum_{n=1}^{M}\gamma_n^{N_2}.$$

Since by assumption $\sum_{n=1}^{\infty}\gamma_n^{N_2}=\infty$, it follows that $\log P_M^{N_2}(\lambda)\to-\infty$, or $P_M^{N_2}(\lambda)\to 0$ as $M\to\infty$. Finally, notice that

$$\prod_{n}^{\infty}(1-\lambda\gamma_n)=\prod_{n=1}^{N_2-1}(1-\lambda\gamma_n)\lim_{M\to\infty}P_M^{N_2}(\lambda)=0,$$

establishing item c.
$(b \Rightarrow a)$. Suppose $\lambda > 0$ is so that $\prod_n (1 - \lambda\gamma_n) = 0$. Choose $N_1 > 0$ so that $\lambda\gamma_n^{N_1} < 1$ for all $n > 0$. Let $\hat{\gamma} = \sup_n \gamma_n^{N_1} < \lambda^{-1}$, and write

$$\log\left(1 - \lambda\gamma_n^{N_1}\right) = -\lambda\gamma_n^{N_1} + \left(\lambda\gamma_n^{N_1}\right)^2 F(\lambda\gamma_n^{N_1}),$$

where F is a continuous function on $[0, \hat{\gamma}]$. Define

$$P_M^{N_1}(\lambda) = \prod_{n=1}^{M} (1 - \lambda\gamma_n^{N_1}),$$

and observe that

$$\log P_M^{N_1}(\lambda) = -\lambda \sum_{n=1}^{M} \gamma_n^{N_1} + \sum_{n=1}^{M} \left(\lambda\gamma_n^{N_1}\right)^2 F(\lambda\gamma_n^{N_1}).$$

Let

$$\hat{F} = \sup_{\gamma \in [0,\hat{\gamma}]} |F(\lambda\gamma)| < \infty.$$

It follows that

$$\sum_{n=1}^{\infty} \left(\lambda\gamma_n^{N_1}\right)^2 |F(\lambda\gamma_n^{N_1})| \le \hat{F}\lambda^2 \sum_{n=1}^{\infty} \left(\gamma_n^{N_1}\right)^2 < \infty,$$

and thus there exists $\delta \in \mathbb{R}$ so that

$$\sum_{n=1}^{M} \left(\lambda\gamma_n^{N_1}\right)^2 F(\lambda\gamma_n^{N_1}) \to \delta \text{ as } M \to \infty.$$

By assumption, $P_M^{N_1}(\lambda) \to 0$ and thus $\log P_M^{N_1}(\lambda) \to -\infty$ as $M \to \infty$. Thus

$$\begin{aligned}
-\infty &= \lim_{M\to\infty} \log P_M^{N_1}(\lambda) = \lim_{M\to\infty} \left(-\lambda \sum_{n=1}^{M} \gamma_n^{N_1} + \sum_{n=1}^{M} \left(\lambda\gamma_n^{N_1}\right)^2 F(\lambda\gamma_n^{N_1}) \right) \\
&= -\lim_{M\to\infty} \lambda \sum_{n=1}^{M} \gamma_n^{N_1} + \lim_{M\to\infty} \sum_{n=1}^{M} \left(\lambda\gamma_n^{N_1}\right)^2 F(\lambda\gamma_n^{N_1}) \\
&= -\lambda \lim_{M\to\infty} \sum_{n=1}^{M} \gamma_n^{N_1} + \delta.
\end{aligned}$$

It follows that

$$\infty = \lim_{M\to\infty} \sum_{n=1}^{M} \gamma_n^{N_1} < \sum_{n=1}^{\infty} \gamma_n,$$

thus establishing item a.

That $(c \Rightarrow b)$ is trivial and the proof is complete. $\blacksquare$

Proof of Theorem 2. Define

$$\overline{Q} = \max\left\{\frac{\phi u(w_{\max})}{1-\beta\alpha\phi}, \frac{u(b)}{1-\beta}\right\} \text{ and } \underline{Q} = \max\left\{\frac{\phi u(w_{\min})}{1-\beta\alpha\phi}, \frac{u(b)}{1-\beta}\right\},$$

where we note that by Assumption B $\underline{Q} < \overline{Q}$. It is clear from equation (8) of Proposition 1 that $\hat{T}(\overline{Q}, \hat{w}_t^N) < \overline{Q}$ and $\hat{T}(\underline{Q}, \hat{w}_t^N) > \underline{Q}$ for all samples $\hat{w}_t^N$. It follows that for any initial Q the sequence is eventually in $[\underline{Q}, \overline{Q}]$. Thus, without loss of generality, we can assume that $Q_0 \in [\underline{Q}, \overline{Q}]$ and therefore that $Q_t \in [\underline{Q}, \overline{Q}]$ for all $t \geq 1$.

From equation (13) we have that

$$Q_{t+1} - Q^* = Q_t - Q^* + \gamma_{t+1}\left(\hat{T}(Q_t, \hat{w}_t^N) - Q_t\right).$$

Denote by $E_t(\cdot)$ the expectations operator conditional on all information available before the time t wage sample is drawn. Observe that

$$\begin{aligned} E_t\left(\hat{T}(Q_t, \hat{w}_t^N)\right) &= N^{-1}\sum_{k=1}^{N} E_t \max\left\{\begin{array}{c} \phi u\left(\hat{w}_t(k)\right) + \beta\alpha\phi Q_t \\ u(b) + \beta Q_t \end{array}\right\} \\ &= N^{-1}\sum_{k=1}^{N} E_t V(\hat{w}_t(k), Q_t) = N^{-1}\sum_{k=1}^{N} T(Q_t) = T(Q_t). \end{aligned}$$

The second equality follows from (7) and the third equality follows from (11) and the random sample assumption. Using this observation we may compute

$$E_t[(Q_{t+1}-Q^*)^2] = (Q_t-Q^*)^2 + 2\gamma_{t+1}(Q_t-Q^*)(T(Q_t)-Q_t) + \gamma_{t+1}^2 E_t\left[\left(\hat{T}(Q_t, \hat{w}_t^N) - Q_t\right)^2\right].$$

As $[\underline{Q}, \overline{Q}]$ is compact and $\hat{T}$ is continuous in Q there exists $M > 0$ such that

$$E_t\left[\left(\hat{T}(Q_t, w_{t+1}) - Q_t\right)^2\right] \leq M$$

for all $Q_t \in [\underline{Q}, \overline{Q}]$.

Note that if $f : [a, b] \to \mathbb{R}$ is continuous and is differentiable everywhere except at a finite number of points $a < x_1 < \cdots < x_n < b$, and, where defined, if $f'(x) < \beta$ then for all $a < x < y < b$ we have that

$$\frac{f(y) - f(x)}{y - x} \leq \beta.$$

To see this, suppose, for example, that $a < x < x_1 < y < x_2$. Then

$$\begin{aligned} \frac{f(y) - f(x)}{y - x} &= \frac{f(y) - f(x_1) + f(x_1) - f(x)}{y - x} \\ &\leq \frac{\beta(y - x_1) + \beta(x_1 - x)}{y - x} = \beta. \end{aligned}$$

The general result is then easily verified.

Applying this observation to T, and using the facts that $T'(Q) \leq \beta$ for all Q except possibly at $Q_{\max}$ and $Q_{\min}$, and that $T(Q^*) = Q^*$, it follows that

$$\frac{T(Q) - Q}{Q - Q^*} \leq \beta - 1$$

for all Q. Define $\lambda = -2(\beta - 1) > 0$. Then

$$\begin{aligned} E_t[(Q_{t+1} - Q^*)^2] &\leq (Q_t - Q^*)^2 + 2\gamma_{t+1}(Q_t - Q^*)(T(Q_t) - Q_t) + \gamma_{t+1}^2 M \\ &\leq \left(1 + 2\gamma_{t+1}\frac{T(Q_t) - Q_t}{Q_t - Q^*}\right)(Q_t - Q^*)^2 + \gamma_{t+1}^2 M \\ &\leq (1 - \lambda\gamma_{t+1})(Q_t - Q^*)^2 + \gamma_{t+1}^2 M. \end{aligned} \tag{31}$$

Following the proof strategy of Bray and Savin (1986), define

$$c_t = (Q_t - Q^*)^2 + \left(\sum_{k=t}^{\infty} \gamma_{t+1}^2\right) M.$$

From Equation (31) we know that c_t is a sub-martingale since

$$\begin{aligned} E_t c_{t+1} &= E_t[(Q_{t+1} - Q^*)^2] + \left(\sum_{k=t+1}^{\infty} \gamma_{t+1}^2\right) M \\ &\leq (1 - \lambda\gamma_{t+1})(Q_t - Q^*)^2 + \gamma_{t+1}^2 M + \left(\sum_{k=t+1}^{\infty} \gamma_{t+1}^2\right) M \\ &\leq (Q_t - Q^*)^2 + \left(\sum_{k=t}^{\infty} \gamma_{t+1}^2\right) M = c_t. \end{aligned}$$

As c_t is bounded from below by 0, we apply the Martingale Convergence Theorem to conclude that c_t converges to some random variable $\tilde{c}$ almost surely. This immediately implies that $(Q_t - Q^*)^2$ converges to some random variable $\tilde{D}$ almost surely. It remains to be shown that $\tilde{D} = 0$ almost everywhere, and thus $Q_t \to Q^*$ almost surely.

Suppose not, then $E(\tilde{D}) > 0$. Convergence almost surely then implies that there exists $L > 0$ and $t^* > 0$ such that $E(Q_t - Q^*)^2 \geq L$ for all $t \geq t^*$. Taking expectations of Equation (31) we have that

$$E[(Q_{t+1} - Q^*)^2] \leq (1 - \lambda\gamma_{t+1})E[(Q_t - Q^*)^2] + \gamma_{t+1}^2 M.$$

Since $\gamma_t \to 0$, we can choose any $N > t^*$ such that $\gamma_{t+1} \leq \frac{L\lambda}{2M}$ for all $t \geq N$. It follows that

$$E[(Q_{t+1} - Q^*)^2] \leq \left(1 - \frac{\lambda}{2}\gamma_{t+1}\right) E[(Q_t - Q^*)^2]$$

for all $t \geq N$. We therefore conclude that

$$E[(Q_t - Q^*)^2] \leq E[(Q_N - Q^*)^2] \prod_{k=N}^{t-1} \left(1 - \frac{\lambda}{2}\gamma_{k+1}\right)$$

for all $t \geq N$. By Lemma A.3, Assumption C implies that $\prod_{k=N}^{\infty} \left(1 - \frac{\lambda}{2}\gamma_{k+1}\right) = 0$ and thus

$$E(\tilde{D}) = \lim_{t\to\infty} E[(Q_t - Q^*)^2] = 0,$$

which is a contradiction. Therefore, we conclude that $Q_t \to Q^*$ almost surely. ■

Appendix B: Proofs of results in Section 4

Computation of $\delta(Q, b, \mu)$. Let

$$\psi = \psi(Q, b, \mu) \equiv F(\bar{w}(Q, b), \mu) + \alpha (1 - F(\bar{w}(Q, b), \mu)),$$

which is the probability of being unemployed at the end of the current period conditional on being unemployed at end of the previous period. Then

$$\begin{aligned} \delta(Q, b, \mu) &= 1 \cdot (1 - \psi) + 2 \cdot \psi \cdot (1 - \psi) + 3 \cdot \psi^2 \cdot (1 - \psi) + \ldots \\ &= (1 - \psi) \sum_{n \geq 0} (n + 1)\psi^n = \frac{1}{(1 - \alpha)(1 - F(\bar{w}(Q, b), \mu))}. \blacksquare \end{aligned}$$

Proof of Lemma 1. We first consider a change to the benefits level b. Implicit differentiation yields $Q_b^* = (1 - DT(Q^*))^{-1} T_b(Q^*) > 0$. As shown in the proof of Lemma A.2, $DT(Q) \in (0, 1)$. Also, since Q^* is in the interior, the T-map is given locally by

$$T(Q) = (u(b) + \beta Q) F(\bar{w}) + \beta\alpha\phi Q(1 - F(\bar{w})) + \phi \int_{\bar{w}(Q)}^{w_{\max}} u(w) dF(w). \tag{32}$$

Direct computation yields

$$\begin{aligned} T_b(Q^*) &= F(w^*)u'(b) + (u(b) + \beta Q^* - \beta\alpha\phi Q^*) dF(w^*) w_b^* - \phi u(w^*) dF(w^*) w_b^* \\ &= F(w^*)u'(b) + [u(b) + \beta Q^* - (\phi u(w^*) + \beta\alpha\phi Q^*)] dF(w^*) w_b^* = F(w^*)u'(b) > 0, \end{aligned}$$

where the term in square brackets equals zero by (9). It follows that $Q_b^* > 0$.

Turning now to the determination of Q_μ^*, observe that we may differentiate (32) to obtain

$$\begin{aligned} T_\mu(Q^*) &= (u(b) + \beta Q^* - \beta\alpha\phi Q^*) \frac{\partial}{\partial\mu} F(w^*) + \frac{\partial}{\partial\mu} \left(\int_{w^*}^{w_{\max}(\mu)} u(w) dF(w, \mu) \right) \\ &= \phi u(w^*) \frac{\partial}{\partial\mu} \left(\int_{w_{\min}(\mu)}^{w^*} dF(w, \mu) + \int_{w^*}^{w_{\max}(\mu)} u(w) dF(w, \mu) \right) \\ &= \phi \frac{\partial}{\partial\mu} \int_{w_{\min}(\mu)}^{w_{\max}(\mu)} \tilde{u}(w) dF(w, \mu), \text{ where } \tilde{u}(w) = \begin{cases} u(w^*) & \text{if} \quad w \leq w^* \\ u(w) & \text{if} \quad w > w^* \end{cases}, \end{aligned}$$

and the second equality exploits (9). A well-known consequence of first-order stochastic dominance (see Exercise 12.9 of Stokey and Lucas (1989)) is that if g is a continuous, non-decreasing function then

$$\mu_1 \leq \mu_2 \implies \int_{w_{\min}(\mu_1)}^{w_{\max}(\mu_1)} g(w)dF(w,\, \mu_1) \leq \int_{w_{\min}(\mu_2)}^{w_{\max}(\mu_2)} g(w)dF(w,\, \mu_2).$$

Since $\tilde{u}$ is a continuous non-decreasing function, it follows by stochastic dominance that $T_\mu > 0$. Since $Q^*_\mu = (1 - DT(Q^*))^{-1}T_\mu$, we conclude $Q^*_\mu > 0$. ■

Proof of Proposition 3. By Lemma 1, we need only establish that $\frac{\partial h}{\partial b} < 0$ and $\frac{\partial h}{\partial Q} < 0$. Since $h = (1-\alpha)(1-F)$, we may compute

$$\frac{\partial h}{\partial b} = -(1-\alpha)dF(w^*)\frac{\partial}{\partial b}\bar{w}\left(Q^*, b\right) < 0 \text{ and } \frac{\partial h}{\partial Q} = -(1-\alpha)dF(w^*)\frac{\partial}{\partial Q}\bar{w}\left(Q^*, b\right) < 0,$$

where the inequalities follow from equation (10). ■

Proof of Corollary 3. First assume that $F(\bar{w}, \mu)$ is given by (18). Direct computation yields

$$F_\mu(\bar{w}, 0) = -\left(\int_{w_{\min}}^{\bar{w}} d^2F(w) + dF\left(w_{\min}\right)\right) = -dF\left(\bar{w}\right).$$

It follows that the condition (17) reduces to $w^*_\mu \leq 1$.

Now assume that $F(\bar{w}, \mu)$ is given by (19). Using the substitution

$$v = (s\sqrt{2})^{-1}(\mu - \log(w)),$$

we may write

$$F(\bar{w}, \mu) = \frac{1}{\sqrt{\pi}}\int_{\frac{\mu - \log(\bar{w})}{s\sqrt{2}}}^{\infty} e^{-v^2}dv,$$

so that

$$F_\mu(\bar{w}, \mu) = -\frac{1}{s\sqrt{2\pi}}e^{-\frac{(\mu - \log(w))^2}{2s^2}}dw,$$

and the result follows from condition (17) and the functional form of $dF(\bar{w}, \mu)$. ■

Proof of Proposition 5. First observe that

$$du_0 = -u^*dh + (1-\alpha)(1-u^*)q_w dw. \tag{33}$$

Next, notice that

$$u^* = \frac{\alpha}{h+\alpha} \implies -u^*dh = \alpha\frac{du^*}{u^*}. \tag{34}$$

If $d\star < 0$ then $dq = q_w dw = 0$. It follows from equations (33)-(34) that $d\star < 0$ implies

$$du_0 = \alpha \frac{du^*}{u^*}.$$

Turning now to the case $d\star \geq 0$, and using the definition of h and that $dq = (1-F)^{-1} dF$, we have that $dh = h_w dw^* + h_\star d\star$ with $h_w = -(1-\alpha)dF$. It follows that

$$h_w dw^* = -(1-\alpha)(1-F)q_w dw^* = -h \cdot q_w dw^*. \tag{35}$$

Combining (33)-(35), we get

$$\begin{aligned}
du_0 &= \alpha \frac{du^*}{u^*} - \frac{(1-\alpha)(1-u^*)}{h}(-h \cdot q_w dw^*) \\
&= \alpha \frac{du^*}{u^*} - \frac{(1-\alpha)(1-u^*)}{h}(dh - h_\star d\star) \\
&= \alpha \frac{du^*}{u^*} - \left(\frac{1-\alpha}{\alpha}\right) u^*(dh - h_\star d\star) \\
&= \alpha \frac{du^*}{u^*} + \left(\frac{1-\alpha}{\alpha}\right)\left(\alpha \frac{du^*}{u^*}\right) + \left(\frac{1-\alpha}{\alpha}\right) u^* h_\star d\star \\
&= \frac{du^*}{u^*} + \left(\frac{1-\alpha}{\alpha}\right) u^* h_\star d\star,
\end{aligned}$$

where the third equality uses that $h^{-1}(1-u^*) = \alpha^{-1}u^*$. ■

References

[1] Adam, K., Marcet, A. and Beutel, J. (2017), Stock price booms and expected capital gains, American Economic Review, 107, 2352-2408.

[2] Bray, M.M. and Savin, N.E. (1986), Rational expectations equilibria, learning, and model specification, Econometrica 54, 1129-1160.

[3] Cogley and Sargent (2008), Anticipated utility and rational expectations as approximations of Bayesian decision making, International Economic Review, 49, 185-221.

[4] Evans, G.W. and Honkapohja, S. (2001), Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.

[5] Evans, G.W. and McGough, B. (2018a), Learning to optimize, mimeo., University of Oregon.

[6] Evans, G.W. and McGough, B. (2018b), Agent-level learning in general equilibrium: the shadow-price approach.

[7] Greene, W.H. (2012), Econometric Analysis, 7th ed. Prentice Hall, Boston MA.

[8] Kreps, D. (1998), Anticipated utility and dynamic choice, in Frontiers of Economic Research: the Nancy L. Schwartz Memorial Lectures, eds. Jacobs, D.P., Kalai, E. and Kamien, M. Cambridge University Press, Cambridge UK.

[9] Krueger, D., Mitman, K. and Perri, F. (2016), Macroeconomics and household heterogeneity, Ch 11 in Handbook of Macroeconomics, Volume 2A, 843-921, eds. Taylor, J.B. and Uhlig, H., Elsevier.

[10] Levy, H. (1973), Stochastic dominance among log-normal prospects, International Economic Review 14(3), 1601-614.

[11] Ljungqvist, L. and Sargent, T.J. (2012), Recursive Macroeconomic Theory, 3rd ed., MIT Press, Cambridge,MA.

[12] Marcet, A. and Sargent, T.J. (1989), Convergence of least squares learning mechanisms in self-referential linear stochastic models, Journal of Economic Theory 48, 337-368.

[13] Marimon, R., McGrattan E. and Sargent, T.J. (1990), Money as a medium of exchange with artificially intelligent agents, Journal of Economic Dynamics and Control 14, 329-373.

[14] McCall, J.J. (1970), Economics of information and job search, Quarterly Journal of Economics 84, 113-126.

[15] Preston, B. (2005), Learning about monetary policy rules when long-horizon expectations matter, International Journal of Central Banking, 1, 81-126.

[16] Stokey, N.L. and Lucas, R. E. (1989), Recursive Methods in Economics, Harvard University Press, Cambridge, MA.

[17] Sutton, R. S. and Barto, A. G. (2011), Reinforcement learning: An introduction, MIT Press, Cambridge, MA.

[18] Watkins, C. (1989), Learning from Delayed Rewards. PhD Thesis, University of Cambridge, Cambridge, UK.

[19] Watkins C. and Dayan, P. (1992), Technical Note: Q-learning," Machine Learning, 8, 279-292.