Matching with Bounded Probabilities: Flexibility, Efficiency, and Crisis Dynamics

Mario Rafael Silva[a,]

[a] *Department of Economics, Northeastern University, 301 Lake Hall, 360 Huntington Avenue, Boston, MA, 02115, USA*

Abstract

We propose the Soft-Min Cobb-Douglas (SMCD) matching function, freely calibratable to a target job-finding rate and elasticity like Cobb-Douglas yet keeping matching probabilities bounded like den Haan–Ramey–Watson (HRW). On U.S. data it dominates HRW and respects the bounds CD violates. We embed the externally estimated two-plateau shape in an equilibrium unemployment model with fixed matching costs, solved globally and calibrated on separate aggregate moments. The shape produces a sign-reversing unemployment gap: the market is inefficiently tight on average yet generates large excess unemployment in crises. The same bounded, cyclically varying elasticity makes unemployment more crisis-prone than under HRW.

Keywords: matching function, probability bounds, unemployment gap, unemployment crises, structural break, simulated method of moments

JEL Classification: J64, E24, E32, C13, C15, C22

1. Introduction

The matching function is a fundamental building block of equilibrium unemployment models in the tradition of Diamond (1982), Mortensen and Pissarides (1994), and Pissarides (1985). It maps unemployment u and vacancies v into a flow of new matches $m(u, v)$, and its properties shape the steady-state and cyclical behavior of any model in which job creation and destruction are mediated by search frictions. Quantitative work typically summarizes it by two numbers at a calibration point: the *level* of matching, captured by a job-finding probability f^*, and the *elasticity* of matching with respect to vacancies, ε^*, which governs how sensitively hiring responds to labor market tightness $\theta = v/u$. Yet the object that matters most for the cyclical behavior of unemployment is the *shape* of the function—how that elasticity varies as tightness moves over the business cycle. We show that this shape has first-order consequences that the standard Cobb-Douglas workhorse cannot even represent:

Email address: `m.silva@northeastern.edu` (Mario Rafael Silva)

it governs both the tail risk of unemployment crises and whether the search equilibrium is constrained-efficient, and it can make a labor market that is inefficiently *tight* on average the source of large excess unemployment in deep recessions.

In practice, a large body of quantitative work relies on the Cobb-Douglas (CD) matching function, $m(u, v) = Au^{\alpha}v^{1-\alpha}$, because it is analytically convenient and empirical surveys find it fits aggregate data reasonably well (Petrongolo and Pissarides, 2001). Its appeal for calibration is also clear: given targets $(f^*, \varepsilon^*, \theta^*)$, the parameters $\alpha = 1 - \varepsilon^*$ and $A = f^*/(\theta^*)^{\varepsilon^*}$ follow immediately from closed-form formulas. Yet the Cobb-Douglas function has a basic deficiency: it does not bound the implied matching probabilities. The job-finding probability $f(\theta) = A\theta^{1-\alpha}$ grows without bound as θ increases, violating $f \leq 1$, and the vacancy-filling probability $q(\theta) = A\theta^{-\alpha}$ fails to respect $q \leq 1$ at low tightness. These violations are empirically relevant. As we document in Section 3, monthly market tightness falls below the lower admissible bound in nearly one in ten months over our 1951–2025 sample, concentrated in every major recession. It is the vacancy-filling bound that fails empirically, not the job-finding bound.

A smooth alternative that satisfies the bounds is the matching function of den Haan et al. (2000), henceforth HRW: $m(u, v) = uv/\left(u^{\nu} + v^{\nu}\right)^{1/\nu}$. The HRW function is smooth, has constant returns to scale, and bounds both f and q in $(0, 1)$ for all θ. However, it introduces a different limitation: the job-finding elasticity $\varepsilon(\theta) = 1/(1 + \theta^{\nu})$ is pinned entirely by the single parameter ν. Given a calibration point θ^*, the researcher can match either the target elasticity ε^* or the target job-finding rate f^*, but not both simultaneously. In practice, den Haan et al. (2000) calibrate ν to match business-cycle moments rather than steady-state targets, acknowledging that joint calibration to level and elasticity is infeasible within this specification. These two functions are the standard benchmarks in the quantitative DMP literature (Petrosky-Nadeau and Wasmer, 2017), and between them they pose a dilemma: Cobb-Douglas is freely calibratable but unbounded, whereas HRW is bounded but too rigid to match level and elasticity jointly.

This paper proposes a matching function, the *Soft-Min Cobb-Douglas* (SMCD), which resolves both limitations:

$$m(u, v) = \frac{sauv}{\left[(av)^p + u^p\right]^{1/p}},$$

with parameters $s \in (0, 1)$, $a > 0$, and $p > 0$. We show that it satisfies four desiderata simultaneously: (i) constant returns to scale; (ii) smoothness and concavity; (iii) exact joint calibration to any target triple $(\theta^*, f^*, \varepsilon^*)$, unlike HRW; and (iv) bounded matching probabilities $f, q \in (0, 1)$ for all θ, unlike CD. In short, the SMCD resolves the dilemma: it is freely calibratable like Cobb-Douglas yet bounded like HRW, which it nests as the special case $s = a = 1$, and near the calibration point it coincides with Cobb-Douglas to first order.

The name *soft-min* arises from the fact that the job-finding rate $f(\theta)$ behaves as a smooth minimum of two limiting regimes: a vacancy-limited regime at low tightness where $f(\theta) \approx sa\theta$ and each vacancy is filled with near certainty, and a worker-limited regime at high tightness where $f(\theta) \approx s$ and workers face a capacity constraint on successful contacts. The parameter p governs how sharply the transition occurs between regimes. As $p \to \infty$, the function converges to the hard-min $\min\{sa\theta, s\}$.

We take this specification to U.S. labor market data spanning 1951–2025, with the aim of letting the data discipline the *shape* of the matching function. Using the job-finding rate constructed by the method of Shimer (2012) and a composite vacancy series that splices Help-Wanted Index data with JOLTS openings (Barnichon, 2010), we estimate SMCD, HRW, and CD and compare their fit. A QLR sup-Wald test with a moving-block bootstrap identifies a structural break in the log-linear relationship between f and θ in October 2008, coinciding with the Lehman Brothers collapse and the onset of the Great Recession. The break reflects both a level shift and a slope change, consistent with the persistent decline in matching efficiency documented by Barnichon and Figura (2015) and Şahin et al. (2014). We let curvature differ between the vacancy-limited and worker-limited regimes through a two-plateau extension that adds a single parameter. Estimated separately on the pre- and post-break subsamples, the two-plateau SMCD strictly improves on the single-p SMCD and, on CD's own admissible tightness range, comes within about two percent of CD pre-break and outperforms it post-break, while always respecting the probability bounds CD violates in about one month in ten. CD's small full-sample edge comes almost entirely from recession months in which it predicts a vacancy-filling probability above one.

The payoff of a data-disciplined shape appears once we embed it in general equilibrium. We place the pre-break estimates in a calibrated Diamond–Mortensen–Pissarides economy and compare the two-plateau SMCD with HRW under a deliberately split identification. We calibrate the shape of the matching function *externally* by finding the curve that best fits data on tightness and job finding rates. The remaining structural parameters—the value of non-work b, the bargaining weight α_L, and the vacancy cost k—are then estimated by simulated method of moments on the globally solved ergodic distribution, targeting a separate, non-overlapping set of aggregate moments: the relative volatility $\sigma(u)/\sigma(z)$, mean tightness $\bar{\theta}$, and the wage elasticity $\varepsilon_{w,y}$. Because these moments discipline levels, volatilities, and wage cyclicality rather than the curvature of matching, the matching-function shape is never a free parameter tuned to fit the business cycle. Any difference across specifications—and the efficiency and crisis results that follow—is therefore a consequence of a data-disciplined matching technology.

To partially address the unemployment volatility puzzle (Shimer, 2005), we adopt the two-part hiring cost of Pissarides (2009): a fixed matching cost κ paid upon formation plus a flow vacancy cost k. This cost structure implies a *modified Hosios condition* (Proposition 3): the efficient bargaining weight becomes $\alpha_L^H(\theta) = (1-\phi)\eta_u(\theta)/[1-\phi\,\eta_u(\theta)]$, where ϕ is the fixed-cost share. The condition generalizes the standard $\alpha_L = \eta_u$ by recognizing that only the flow component creates a congestion externality. Under Cobb-Douglas, η_u is constant, so the modification is state-dependent only through $\phi(\theta)$—a mild variation. Under a bounded matching function whose elasticity $\eta_u(\theta)$ varies with the cycle, the condition acquires quantitative bite: in recessions, $\eta_u(\theta) \to 0$ drives $\alpha_L^H \to 0$, and a constant bargaining weight that is nearly efficient on average becomes grossly excessive.

Two results follow. First, and more importantly, the state-dependent Hosios condition produces a *sign-reversing unemployment gap*. On average the decentralized 2P-SMCD economy is inefficiently tight—firms over-hire, so unemployment sits below its efficient level, consistent with the evidence of Michaillat and Saez (2024). However, in crises, the matching elasticity collapses and the constant bargaining weight overshoots the vanishing efficient share, and the gap flips to large *excess* unemployment. HRW, whose calibrated weight sits near the modified Hosios boundary throughout, exhibits no such sign flip. Second, and reinforcing the first, the near-Leontief recession branch generates vacancy shutdowns and fat-tailed crisis dynamics: the simulated unemployment distribution is an order of magnitude more right-skewed under the two-plateau SMCD than under HRW (skewness 10.7 versus 1.8). This is the Petrosky-Nadeau and Zhang (2017, 2021) crisis channel, which we show sharpens under a matching function jointly fitting level and elasticity targets. Both results emerge only from the global solution of a model with a bounded, cyclically-varying matching elasticity.

The paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the data and documents the CD bound violations. Section 4 develops the SMCD matching function and proves its properties, with a numerical example calibrated to the sample means from Section 3. Section 5 presents the empirical horse race. Section 6 embeds the estimated matching functions in a calibrated DMP model and shows that the shape of the matching function produces a state-dependent efficiency wedge with a sign-reversing unemployment gap and governs the tail risk of unemployment crises. Section 7 addresses the objection that probability bounds are irrelevant in continuous-time or high-frequency formulations. Section 8 concludes. The appendices provide additional figures, proofs and derivations, details of the global solution method and estimation, and robustness checks.

2. Related Literature

Petrongolo and Pissarides (2001) provide the canonical survey of matching function estimation. They find that constant returns to scale is broadly consistent with the data and that CD fits aggregate time-series data reasonably well, with estimated unemployment elasticities in the range $[0.5, 0.7]$. Their survey, however, focuses on the interior of the data and does not address the behavior of CD at extreme tightness values where probability bounds are violated. Our paper takes this limitation as its starting point.

The HRW matching function, introduced by den Haan et al. (2000), bounds matching probabilities by construction and has been adopted by Hagedorn and Manovskii (2008) and Petrosky-Nadeau and Zhang (2017), who show that an accurate global solution of the DMP model is critical for capturing the nonlinear dynamics that arise as matching probabilities approach their bounds in recessions. However, HRW's single-parameter structure prevents joint calibration to both level and elasticity targets. The SMCD nests HRW as a special case ($s = a = 1$) while adding the flexibility needed for joint calibration.[1] The sensitivity of quantitative predictions to matching function parameters in Shimer (2005), Hagedorn and Manovskii (2008), and Petrosky-Nadeau and Zhang (2017) underscores the importance of being able to hit calibration targets exactly.[2]

Other approaches include Anderson and Burgess (2000), who reject CD's homogeneity restrictions, and Borowczyk-Martins et al. (2013), who decompose matching into search intensity and contact technology. Stevens (2007) and Angelis and Bramoullé (2026) develop structural microfoundations for the CES family. The present paper instead asks which reduced-form specification best fits aggregate data while respecting the probability bounds that bind in recessions.[3] Lange and Papageorgiou (2020) nonparametrically estimate the matching function and find that vacancy elasticity is higher at low tightness—consistent with the countercyclical elasticity that probability bounds mechanically require.

The empirical literature has documented substantial time variation in U.S. matching efficiency. Barnichon and Figura (2015) find a sharp decline beginning in 2008–2009, which they attribute to compositional shifts and mismatch. Şahin et al. (2014) estimate that

[1]The urn-ball matching function, $q(\theta) = \left(1 - e^{-1/\theta}\right)$, $f(\theta) = \theta\, q(\theta)$, shares the same one-parameter limitation and is strictly less flexible than HRW in shape. A scaled HRW, $f(\theta) = A\theta/(1 + \theta^{\nu})^{1/\nu}$ with $A \leq 1$, would allow joint calibration—but it is literally the SMCD at $s = A$, $a = 1$. Adding an asymmetric weighting parameter recovers the full SMCD parameterization.

[2]The same functional-form requirements arise in New Monetarist models of monetary exchange (Kiyotaki and Wright, 1993; Lagos and Wright, 2005) and in goods-market models of capacity utilization (Bai et al., 2025; Silva and Urias, 2025), where meeting probabilities must satisfy the same bounds. The SMCD is equally applicable in these settings.

[3]Angelis and Bramoullé (2026) show that a dense-network microfoundation requires $\nu \leq 1$. HRW's original $\hat{\nu} = 1.27$ falls outside this range.

mismatch accounts for a substantial fraction of the post-2008 unemployment increase. Our structural break finding at October 2008 aligns with this evidence.

A separate strand studies the *efficiency* of the search equilibrium. Michaillat and Saez (2021, 2024) develop a sufficient-statistic approach that measures the unemployment gap $u - u^*$—and the full-employment rate from the elasticity of the Beveridge curve and the relative social costs of unemployment and vacancies, without committing to a particular matching function or wage-setting protocol. Landais et al. (2018) build on the same framework to study optimal unemployment insurance. That model-free robustness rests on an isoelastic Beveridge curve and symmetric recruiting costs. Section 6 takes a complementary, structural route. We embed the two-plateau matching function in an equilibrium unemployment model with two-part hiring costs (Pissarides, 2009), characterize efficiency through a modified Hosios condition, and trace the resulting unemployment gap over the cycle.

3. Data

3.1. Sources and construction

We use monthly U.S. data from January 1951 through December 2025. The unemployment rate is the civilian unemployment rate (UNRATE) from FRED, expressed as a fraction of the labor force. Vacancy data splice two series. The Barnichon (2010) composite Help-Wanted Index (HWI), which adjusts newspaper help-wanted counts for the secular shift to online job postings and is expressed as vacancies per labor force, runs from 1951 through 2019. JOLTS job openings (JTSJOL), normalized by the civilian labor force (CLF16OV), begin in December 2000. Both series are available over the December 2000 – December 2019 overlap window. We use this overlap only to estimate the level rescaling of the HWI onto the JOLTS scale via a log-mean ratio, and then use JOLTS itself from December 2000 onward. The spliced series is thus the scaled HWI through November 2000 and JOLTS thereafter. Market tightness is $\theta_t = v_t/u_t$. Monthly job-finding and separation rates follow the continuous-time construction of Shimer (2012), using unemployment levels (UNEMPLOY), employment levels (CE16OV), and short-term unemployment (UEMPLT5).

3.2. Summary statistics and time series

Figure 1 plots the unemployment rate and vacancy rate (left panel) and market tightness (right panel) over the full sample. Periods in which $\theta < \theta_{\text{low}} \approx 0.34$—where the Cobb-Douglas matching function implies $q > 1$—are marked in red. These violations are confined to recessions and account for 13.5% of all months in the sample.

Figure 1: Monthly U.S. labor market data, 1951–2025. Top: unemployment rate and vacancy rate (Barnichon 2010 composite HWI spliced with JOLTS). Bottom: market tightness $\theta = v/u$. Red dots mark months with $\theta < \theta_{\text{low}}$ (dashed line), where the Cobb–Douglas matching function implies $q > 1$. Grey bands are NBER recession dates.

Table 1 reports sample means over the full period.[4]

3.3. *Preliminary evidence on the matching function*

The matching function $m(u, v)$ maps unemployment u and vacancies v into a flow of new matches. Under constant returns to scale, the job-finding probability depends only on market tightness $\theta = v/u$:

$$f(\theta) \equiv \frac{m(u, v)}{u} = m(1, \theta).$$

[4]The sample mean of $\theta_t = v_t/u_t$ (0.70) exceeds the ratio of sample means $\bar{v}/\bar{u} = 0.61$ due to Jensen's inequality and the positive covariance between vacancies and the inverse of unemployment over the cycle. We use $\bar{\theta} = 0.70$ as the calibration target, consistent with treating θ^* as the mean of the tightness series rather than the ratio of its component means.

Table 1: Sample means, U.S. monthly data 1951–2025.

Variable	Mean	Std. dev.	Min	Max
Unemployment rate u	0.057	0.017	0.025	0.148
Vacancy rate v	0.035	0.010	0.014	0.075
Market tightness $\theta = v/u$	0.700	0.358	0.152	2.024
Job-finding rate f	0.402	0.096	0.162	0.691

Note: Monthly, Jan 1951 – Dec 2025. $n = 896$. All rates are fractions of the labor force. Vacancy series splices Barnichon (2010) HWI with JOLTS.

The matching balance condition $m_t = u_t f_t = v_t q_t$ then pins down the vacancy-filling probability as $q_t = f_t/\theta_t$.

We measure f_t following Shimer (2012). Let $u_t^{<5}$ denote the stock of workers unemployed for less than five weeks (BLS series UEMPLT5), which proxies the flow of new job-losers entering unemployment during month t. Write next-period unemployment as $u_{t+1} = u_t(1 - f_t) + u_{t+1}^{<5}$ and rearrange for the job finding rate $f_t = 1 - (u_{t+1} - u_{t+1}^{<5})/u_t$.

We estimate the Cobb-Douglas matching function on the full sample by nonlinear least squares (NLS) in levels, $f_t = A\theta_t^{\varepsilon} + e_t$, the same estimator used throughout the horse race of Section 5. This yields $\hat{A} \approx 0.45$ and $\hat{\varepsilon} \approx 0.26$ ($\hat{\alpha} = 1 - \hat{\varepsilon} \approx 0.74$), consistent with Shimer (2005) and within the range surveyed by Petrongolo and Pissarides (2001). These estimates imply the CD probability bounds hold only for $\theta \in [0.34,\ 22.6]$, with 13.5% of monthly observations violating the lower bound. Section 5 examines whether SMCD fits these data better than CD or HRW, and whether the relationship between f and θ is stable over the sample.

4. The Soft-Min Cobb-Douglas Matching Function

We propose the Soft-Min Cobb-Douglas (SMCD) matching function, defined as

$$m(u, v) = \frac{sauv}{\left[(av)^p + u^p\right]^{1/p}}$$

with parameters $s \in (0, 1), a > 0, p > 0$. The denominator is the L^p norm of the vector (u, av). The HRW matching function is a special case with $s = a = 1$. The matching probabilities are

$$f(\theta) = \frac{m}{u} = \frac{sa\theta}{\left[(a\theta)^p + 1\right]^{1/p}} = \frac{s}{(1 + (a\theta)^{-p})^{1/p}} \tag{1}$$

$$q(\theta) = \frac{m}{v} = \frac{sa}{[(a\theta)^p + 1]^{1/p}}$$

The elasticity of the job finding rate is

$$\varepsilon(\theta) = \frac{(a\theta)^{-p}}{1 + (a\theta)^{-p}} = \frac{1}{(a\theta)^p + 1} \tag{2}$$

Note that $f(\theta) \leq s$ and $q(\theta) \leq sa$ for all θ. Moreover, $\varepsilon(0) = 1$ and decreases monotonically to zero as $\theta \to \infty$. This countercyclical pattern is mechanically implied by the bound: since $f(\theta) \leq s$, any bounded increasing job-finding rate must saturate as $\theta \to \infty$, driving $\varepsilon(\theta) \to 0$ in tight markets. The point $a\theta = 1$ characterizes the cross-over at which $\varepsilon = 1/2$.

Constant returns to scale is immediate: $m(\lambda u, \lambda v) = \lambda^2 sauv/\{\lambda[(av)^p + u^p]^{1/p}\} = \lambda m(u, v)$.

Lemma 1. *Suppose $m(u, v)$ has constant returns to scale. Then m is concave if and only if $f(\theta) \equiv m(1, \theta)$ is concave.*

The proof is standard: sufficiency follows from the restriction of a concave function to an affine set. Necessity follows from the convex-combination decomposition $m(tu_1 + (1-t)u_2,\ tv_1 + (1-t)v_2) \geq t\, m(u_1, v_1) + (1-t)\, m(u_2, v_2)$ using the weights $\lambda_i = tu_i/u'$. It therefore suffices to verify concavity of f.

Lemma 2. *The job finding probability $f(\theta) = sa\theta[1 + (a\theta)^p]^{-1/p}$ is concave for all $a, p > 0$.*

Proof. See Appendix B.1. □

Let us check (iii) (joint calibration). That is, given θ^*, we require the matching function to satisfy any $f^* = f(\theta^*) \leq \theta^*$ and elasticity ε^*. First, given p, we characterize choices of a and s which satisfy the job-finding rate and elasticity. Then we find a value of p such that the bounds on the matching probabilities are satisfied. As an intermediate step, we write the job-finding rate in terms of the elasticity using (1) and (2):

$$f(\theta) = s(1 - \varepsilon(\theta))^{1/p}$$

This interpretation emphasizes the fact that $f(\theta)$ is bounded above by s, and that the deviation depends on the elasticity scaled by $1/p$. Given p, choosing $a = (\theta^*)^{-1}[1/\varepsilon^* - 1]^{1/p}$ and $s = f^*/(1 - \varepsilon^*)^{1/p}$ ensures $\varepsilon(\theta^*) = \varepsilon^*$ and $f(\theta^*) = f^*$. To bound the matching probabilities, we need p such that $s \leq 1$ and $sa \leq 1$:

$$s = f^*/(1 - \varepsilon^*)^{1/p} \leq 1, \quad sa = (f^*/\theta^*)(\varepsilon^*)^{-1/p} \leq 1 \tag{3}$$

We can always satisfy these requirements for p sufficiently large since for $p \to \infty$ these requirements reduce to simply $f^* \leq 1$ and $f^* \leq \theta^*$, which is simply a requirement that the approximation target make sense. Specifically, it follows from (3) that it is sufficient to take $p \geq p_{\min}$ where

$$p_{min} = \max\{p_s, p_{sa}\}, \quad p_s = \log(1-\varepsilon^*)/\log(f^*), \quad p_{sa} = \log(\varepsilon^*)/\log(f^*/\theta^*)$$

We thus conclude that the SMCD matching function satisfies all properties (i)–(iv).

The SMCD matching function is isomorphic to a CES matching function $m(u, v) = A[\eta\, u^\rho + (1-\eta)\, v^\rho]^{1/\rho}$. Applying the identity $xy/(x^p + y^p)^{1/p} = (x^{-p} + y^{-p})^{-1/p}$ with $x = u$, $y = av$ rewrites SMCD as $m = s[u^{-p} + a^{-p}v^{-p}]^{-1/p}$. This form matches CES given the mapping

$$\rho = -p, \qquad \eta = \frac{a^p}{1+a^p}, \qquad A = \frac{sa}{(1+a^p)^{1/p}},$$

with inverse $p = -\rho$, $a = [\eta/(1-\eta)]^{-1/\rho}$, $s = A\eta^{1/\rho}$.

Despite this formal equivalence, the SMCD parameterization is preferable for three reasons. First, satisfying the probability bounds requires $\rho < 0$ in the CES, a less obvious restriction than $p \geq p_{\min}$. Second, the SMCD bound conditions $s \leq 1$ and $sa \leq 1$ are immediately legible from (1) and yield $p_{\min}$ in two lines, whereas the equivalent CES constraints $A \leq \eta^{-1/\rho}$ and $A \leq (1-\eta)^{-1/\rho}$ are opaque joint restrictions that give no guidance on how to choose ρ. Third, the CES share $\eta = a^p/(1+a^p)$ conflates the crossover tightness and the curvature without a direct economic interpretation. By contrast, the SMCD parameters s, a, and p each carry a transparent meaning (upper bound, inverse crossover, bottleneck sharpness). HRW is itself CES with $\eta = \frac{1}{2}$, $\rho = -\nu$, $A = 2^{-1/\nu}$, but features a more transparent parameterization. SMCD relaxes HRW's symmetry restriction ($s = a = 1$) in the minimal way needed for joint calibration to level and elasticity targets.

The job-finding rate can be rewritten as

$$f(\theta) = \frac{1}{(s^{-p} + (sa\theta)^{-p})^{1/p}},$$

the generalized soft-min of two caps: a worker-limited ceiling s and a vacancy-limited ceiling $sa\theta$. As $p \to \infty$, $f \to \min\{sa\theta, s\}$. At $p = 1$ it reduces to the harmonic mean $sa\theta/(1 + a\theta)$, which at $s = a = 1$ is the Kiyotaki and Wright (1993) bilateral random meeting function. At low tightness the vacancy-limited regime binds ($f \approx sa\theta$, $\varepsilon \approx 1$). At high tightness the worker-limited regime binds ($f \approx s$, $\varepsilon \approx 0$). The crossover $\theta = 1/a$ marks $\varepsilon = 1/2$. p governs the sharpness of the transition. Figure 2 illustrates both regimes at the baseline calibration $(\theta^*, f^*, \varepsilon^*) = (0.70, 0.40, 0.29)$ with $p = p_{\min}$.

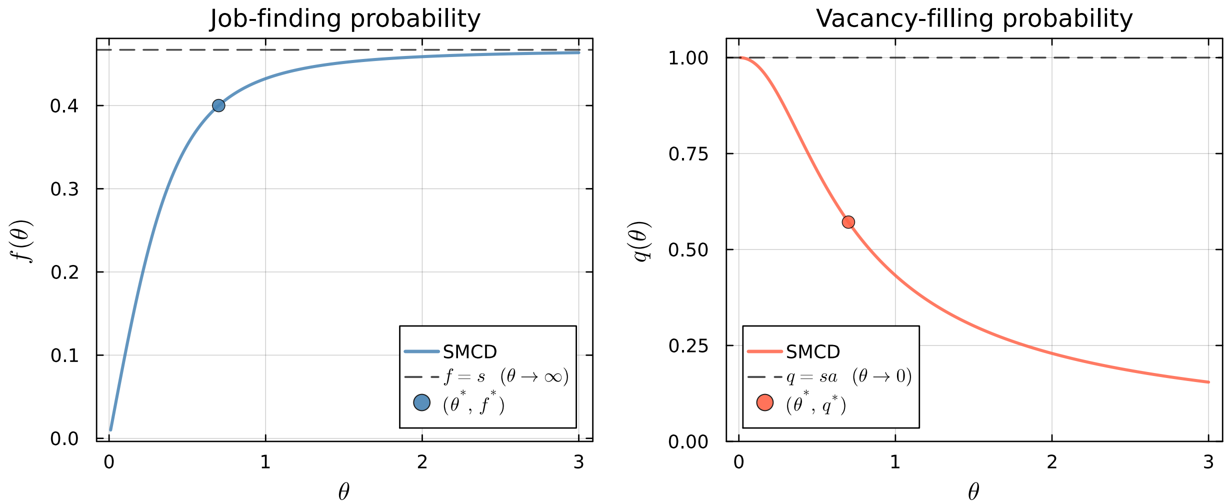

Figure 2: Job-finding probability $f(\theta)$ (left) and vacancy-filling probability $q(\theta)$ (right) under SMCD. Calibration: $(\theta^*, f^*, \varepsilon^*) = (0.70, 0.40, 0.29)$, $p = p_{\min} \approx 2.21$. Dashed lines: asymptotic bounds $f = s$ and $q = sa$. Filled circle: calibration point (θ^*, f^*).

Local Cobb-Douglas behavior is a corollary of (ii) and (iii): smoothness and exact calibration give $d \log m = (1 - \varepsilon^*) d \log u + \varepsilon^* \, d \log v$ at θ^*. As a numerical illustration, we calibrate to U.S. sample means $(\theta^*, f^*, \varepsilon^*) = (0.70, 0.40, 0.29)$ from Section 3. These targets imply $p_{\min} \approx 2.21$, with the vacancy-filling constraint $sa \leq 1$ binding. Setting $p = p_{\min}$ yields $s \approx 0.467$ and $a \approx 2.14$, so that $sa \approx 1.000$.

At this calibration, CD implies $q > 1$ for $\theta < 0.34$—13.5% of monthly observations (Section 3). HRW respects both bounds but its elasticity at θ^* differs from ε^* by construction. No other specification in the quantitative DMP literature satisfies all four properties.

The bound constraint pins p only from below: any $p \geq p_{\min}$ is feasible. Since CD has zero curvature in log space, the SMCD closest to CD minimizes $|\kappa(\theta^*)| = p \, \varepsilon^*(1 - \varepsilon^*)$, which is increasing in p.

Proposition 1. *The SMCD that is closest to Cobb-Douglas in the neighborhood of θ^*, subject to $p \geq p_{\min}$, has $p = p_{\min}$.*[5]

As p rises from $p_{\min}$ toward infinity, the job-finding rate approaches the Leontief profile $\min\{sa\theta, s\}$ (Figure A.2 in the appendix). Figure 3 plots the matching elasticity $\varepsilon(\theta)$ and its

[5]The log approximation $\log f(\theta) \approx \log f^* + \varepsilon^* \Delta - \frac{1}{2} p \, \varepsilon^*(1 - \varepsilon^*)\Delta^2$, with $\Delta = \log(\theta/\theta^*)$, shows that the second-order departure from Cobb-Douglas is proportional to p. Setting $p = p_{\min}$ minimizes this departure while satisfying the probability bounds.

log-curvature $\kappa(\theta) = -p\,\varepsilon(1-\varepsilon)$ at $p = p_{\min}$. The elasticity is strictly decreasing in θ, passing through $\varepsilon^* = 0.29$ at θ^* and through $1/2$ at the crossover $\theta = 1/a \approx 0.47$. $|\kappa|$ is maximized at the crossover. This non-constant elasticity distinguishes SMCD from Cobb-Douglas and aligns with Lange and Papageorgiou (2020).

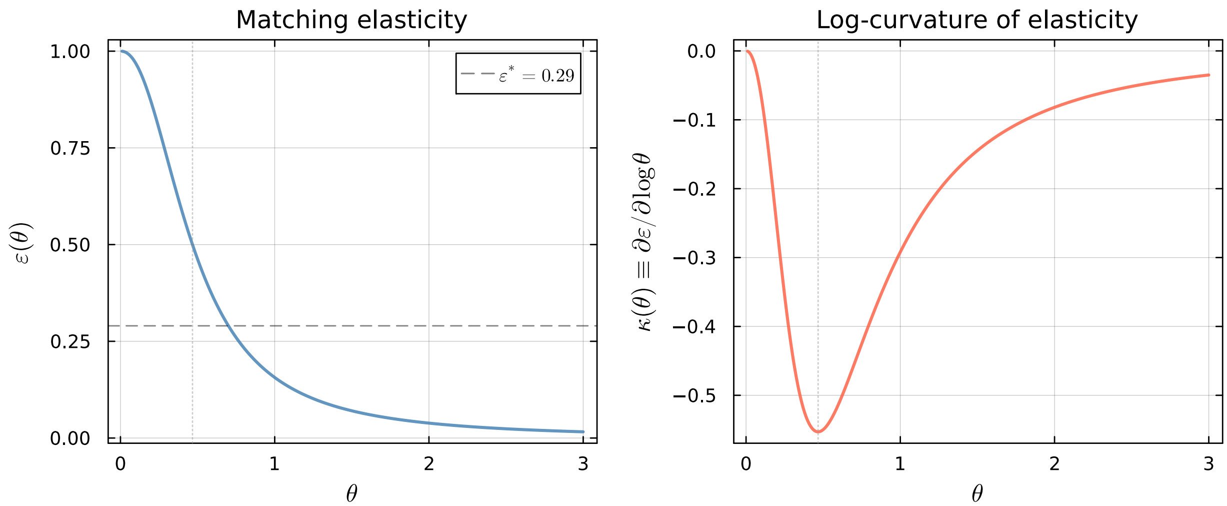

Figure 3: Matching elasticity $\varepsilon(\theta)$ (left) and log-curvature $\kappa(\theta) = \partial\varepsilon/\partial\log\theta$ (right) for the SMCD at $p = p_{\min} \approx 2.21$. Dashed horizontal line: $\varepsilon^* = 0.29$. Vertical dotted line: crossover $\theta = 1/a \approx 0.47$.

4.1. Search Efficiency and the Hosios Condition

Under constant returns to scale, the unemployment elasticity of the matching function is $\eta_u(\theta) \equiv \partial \ln m/\partial \ln u = 1 - \varepsilon(\theta)$. In the standard DMP model with purely flow vacancy costs, Hosios (1990) shows that the decentralized equilibrium is constrained efficient if and only if the worker's bargaining power equals this elasticity, $\alpha_L = \eta_u(\theta) = 1 - \varepsilon(\theta)$. When $\alpha_L > \eta_u$, workers extract more surplus than their matching contribution warrants, wages are too high, and equilibrium unemployment exceeds the efficient level.

Under the SMCD, η_u is not a constant. Combining (2) with $\eta_u(\theta) \equiv 1 - \varepsilon(\theta)$ gives

$$\eta_u(\theta) = \frac{(a\theta)^p}{1 + (a\theta)^p},$$

which is strictly increasing in θ, ranging from zero as $\theta \to 0$ to one as $\theta \to \infty$. The Hosios condition is therefore state-dependent: calibrating $\alpha_L = \eta_u(\theta^*)$ at the steady state, the economy has $\alpha_L > \eta_u(\theta)$ in recessions (low θ), so workers' bargaining power exceeds their matching contribution in states where unemployment is already elevated. This countercyclical wedge is a structural consequence of bounded probabilities. Section 6.2 generalizes the condition to account for fixed matching costs and quantifies the resulting unemployment gap.

Cobb-Douglas suppresses this mechanism entirely: $\eta_u^{CD} = 1 - \varepsilon_{CD}$ is constant, so Hosios holds uniformly or fails uniformly regardless of the cycle. HRW qualitatively replicates the countercyclical pattern but mismeasures its level. At the pre-break sample mean $\bar{\theta} = 0.67$, HRW implies $\eta_u^{HRW} \approx 0.39$ while SMCD implies $\eta_u \approx 0.60$—a gap of 0.21 in the efficient bargaining power. Any welfare assessment built on HRW's η_u is misspecified by construction.

5. Empirical Application

5.1. Structural break in the log-linear matching relationship

The preliminary Cobb-Douglas estimate in Section 3 pools seventy-four years of data that span large institutional changes in how jobs are posted and filled. We test whether the log-linear relationship between f_t and θ_t is stable over the full sample using a QLR sup-Wald test for an unknown structural break. Whereas a standard Wald test requires a pre-specified break date, the QLR procedure computes a Wald statistic at every candidate date in a trimmed interior of the sample and takes the supremum. Because maximizing over candidate dates inflates the test statistic relative to a single-date Wald test, inference relies on the nonstandard asymptotic distribution of Andrews (1993), which we implement via a moving-block bootstrap.

Procedure. Let $y_t = \log f_t$ and $x_t = \log \theta_t$. We fit $y_t = \alpha + \beta x_t + u_t$ on the full sample and test for a break in (α, β) at an unknown date τ. For each candidate τ in a trimmed grid (15% from each end, yielding 629 candidates), we estimate the regression separately on $[1, \tau]$ and $[\tau + 1, n]$, compute Newey–West HAC covariance matrices with bandwidth $L = 7$, and form the Wald statistic $W(\tau) = \Delta(\tau)'\big[V_1(\tau) + V_2(\tau)\big]^{-1}\Delta(\tau)$ where $\Delta(\tau) = b_1(\tau) - b_2(\tau)$. The QLR statistic $\sup_\tau W(\tau)$ has a nonstandard asymptotic distribution under an unknown break date. We obtain p-values from a moving block bootstrap (MBB, block length 14, $B = 499$ draws) applied to centered residuals from the null model.

Results. Figure A.3, reported in the appendix, plots $W(\tau)$ over time. The sup-Wald statistic is $\sup W = 745.97$ with a bootstrap p-value of 0.000. The estimated break date is October 2008, immediately following the September 2008 collapse of Lehman Brothers and the onset of the global financial crisis.[6]

[6] The QLR sup-Wald test isolates the single most important break at an unknown date. A multiple-break procedure such as Bai and Perron (1998) would instead partition the seventy-four-year sample into several regimes. We favor the single-break test on grounds of parsimony: the October 2008 break is overwhelmingly dominant ($\sup W = 745.97$) and is the regime change tied to the matching-efficiency decline that motivates our pre/post split. Whatever additional breaks a multiple-change procedure might detect are far smaller in magnitude and immaterial for the horse race.

Table 2 reports the OLS estimates on each subsample. The intercept shifts from -0.691 to -1.144, a change of -0.453 in logs. At any given tightness, the job-finding rate post-October 2008 is approximately $e^{-0.453} \approx 0.64$ times its pre-2008 level—a persistent 36% decline in matching efficiency. The tightness elasticity $\hat{\varepsilon}$ falls modestly from 0.318 to 0.283, indicating that hiring became slightly less responsive to vacancy creation after the crisis.

Table 2: Estimated log-linear matching relationship, pre- and post-October 2008.

	Pre-break (Jan 1951 – Sep 2008)	Post-break (Oct 2008 – Dec 2025)
Intercept $\hat{\alpha}$	-0.691	-1.144
Tightness elasticity $\hat{\varepsilon}$	0.318	0.283
Implied level $\hat{A} = e^{\hat{\alpha}}$	0.501	0.319
Unemployment elasticity $\hat{\eta}_u = 1 - \hat{\varepsilon}$	0.682	0.717
CD lower bound $\hat{A}^{1/\hat{\eta}_u}$	0.36	0.20

Note: Newey–West HAC, $L = 7$. Break identified by QLR sup-Wald with MBB p-values. This table's log-linear OLS is the standard form required by the break test. The post-break CD lower bound therefore differs slightly (0.20 here versus 0.21 in Tables 3–5) from the NLS-in-levels estimate used throughout the horse race.

Table 2 reveals that the structural break is overwhelmingly a *level shift*: $\hat{A}$ drops from 0.501 to 0.319 (a 36% decline) while $\hat{\varepsilon}$ falls only from 0.318 to 0.283. In the DMP framework, a decline in A shifts the Beveridge curve outward without rotating it, because slope depends on the matching elasticity and separation rate, both essentially unchanged. This is consistent with Lubik (2013), who shows the post-2008 outward movement reflects a structural decline in match efficiency, and with Elsby et al. (2015), who identify lateral Beveridge-curve shifts as the dominant source of unemployment-vacancy dynamics. The post-GFC matching efficiency decline has been attributed to compositional shifts and mismatch (Barnichon and Figura, 2015; Şahin et al., 2014). Lange and Papageorgiou (2020) show that the decline in hires was driven almost entirely by collapsing matching efficiency rather than by changes in search intensity.

5.2. Horse race: SMCD vs. Cobb-Douglas vs. HRW

We now estimate all three matching functions separately on the pre-break (January 1951 – September 2008) and post-break (October 2008 – December 2025) subsamples and compare fit. The pooled full-sample horse race is reported for completeness in Appendix A.

All three models are estimated by nonlinear least squares (NLS), minimizing $\sum_{t=1}^{T}\Big[f_t - \hat{f}(\theta_t;\, \boldsymbol{\psi})\Big]^2$ in levels of the job-finding rate. Estimating all models against the same objective ensures a fair comparison of fit. SMCD has three free parameters $\boldsymbol{\psi} = (s, a, p)$, estimated subject to the probability-bound constraints $s \leq 1$ and $sa \leq 1$. HRW has one unconstrained parameter (ν). Cobb-Douglas has two parameters (A, α) and, for consistency, is estimated by NLS in levels without any probability constraint.[7] Fit is measured in levels of f by root mean square error (RMSE), mean absolute error (MAE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The last two measures penalize the additional parameters of SMCD relative to CD and HRW. For Cobb-Douglas, the NLS and log-OLS estimates are close: the largest deviation is $\Delta\hat{\varepsilon} = 0.019$ in the pre-break subsample, with negligible differences post-break, and the implied $q > 1$ threshold $\hat{\theta}_{\min}$ is unchanged.[8]

Cobb-Douglas constitutes a valid probability model only on a bounded range of tightness. The estimated function $f(\theta) = A\theta^{\varepsilon}$ satisfies $q(\theta) = A\theta^{\varepsilon-1} \leq 1$ only for $\theta \geq \theta_{\min} \equiv A^{1/(1-\varepsilon)}$, and $f(\theta) \leq 1$ only for $\theta \leq \theta_{\max} \equiv A^{-1/\varepsilon}$. Evaluated at the CD estimates of Table 4, $(\theta_{\min}, \theta_{\max}) = (0.36,\ 7.50)$ pre-break and $(0.21,\ 68.1)$ post-break. The upper bound is never reached in either subsample, so the binding constraint is always $\theta_{\min}$: CD violates $q \leq 1$ in 12.0% of pre-break and 9.4% of post-break months, concentrated in recessions. The threshold $\theta_{\min}$ shifts left after the Lehman collapse because the decline in matching efficiency A widens the range of tightness over which CD overstates the vacancy-filling rate. To gauge how much of CD's fit advantage depends on these bound-violating months, Table 3 reports a restricted RMSE* computed on the admissible domain $\theta > \theta_{\min}$, where every specification is a valid probability model.

[7]The log-OLS solution—which minimizes the sum of squared log residuals—serves as the starting point for the NLS routine and is not itself reported as a final estimate.

[8]Because f_t and θ_t are both built from labor-market stocks, they share measurement error in u_t and are jointly determined. Two facts mitigate the concern: NLS is consistent when θ_t is predetermined with respect to the matching shock, as the slow-moving unemployment state is. The upward bias from time-varying matching efficiency (Lange and Papageorgiou, 2020) affects all three specifications equally, leaving the relative rankings intact. The exercise is a calibration in the tradition of Petrosky-Nadeau and Zhang (2017), not a causal identification.

Table 3: Matching function fit statistics, pre- and post-October 2008.

Period	Model	Params	RMSE*	RMSE	MAE	AIC	BIC	q>1 (%)
Pre-2008	SMCD	3	0.0503	0.0564	0.0448	−3978.8	−3965.2	0.0
	CD	2	0.0447	0.0442	0.0349	−4318.0	−4308.9	12.0
	HRW	1	0.0578	0.0674	0.0549	−3735.9	−3731.4	0.0
Post-2008	SMCD	3	0.0296	0.0313	0.0238	−1400.1	−1390.2	0.0
	CD	2	0.0293	0.0284	0.0224	−1441.5	−1434.8	9.4
	HRW	1	0.0532	0.0573	0.0482	−1158.8	−1155.4	0.0

Note: $n = 693$ pre-break, $n = 203$ post-break. RMSE* is the restricted RMSE computed on months with $\theta > \theta_{\min}$ ($\theta_{\min} = 0.36$ pre-break, 0.21 post-break), the domain on which CD also satisfies $q \leq 1$. q>1 % is the fraction of months where the estimated model implies $q > 1$. It is zero by construction for SMCD and HRW.

The post-break data lie uniformly below the pre-break data at any given θ, consistent with the 36% decline in matching efficiency identified in Table 2. Figure 6 in Section 5.4 plots the fitted job-finding rate against the data for every specification, pre- and post-break.

Table 3 delivers three findings. First, on the admissible domain where every specification respects the probability bounds, the restricted RMSE* shows SMCD's cost of guaranteeing validity everywhere is small: 12% above CD pre-break and only 1% above post-break. CD wins the unrestricted full-sample RMSE by wider margins (27% and 10%) and posts the best information criteria, but only because it is estimated without probability constraints and lowers its residuals by extrapolating into the region where $q > 1$ in recession months. The like-for-like comparison is RMSE*, and on it the two specifications are close. The pre-break gap narrows after 2008 because the post-break data span a wider range of tightness (from COVID troughs near $\theta = 0.1$ to the post-pandemic boom above $\theta = 2$), and SMCD handles this dispersion better than a pure power law.

Second, the $q > 1$ and RMSE* columns together show that CD's advantage lives in the months where it violates the bound. In the pre-break subsample, CD predicts a vacancy-filling probability exceeding one in 12.0% of months. Post-break, it is 9.4%. In those months, CD achieves a lower residual for f while simultaneously predicting an impossible q. This modest price is closed entirely by the two-plateau extension 5.4. The natural alternative a practitioner might consider—censoring CD by capping q at one—introduces a C^0 kink at the censoring threshold that precludes the smooth global solution of the Bellman equation on which the quantitative exercise of Section 6 relies.

Third, HRW fits substantially worse than both CD and SMCD (RMSE ratios of 1.19 and 1.83). This holds HRW to a standard it was not built for. den Haan et al. (2000) calibrated HRW's single parameter to business-cycle moments, not jointly to a level and an elasticity target. The horse race asks whether one parameter can span both margins at once, and it cannot. Despite satisfying the probability bounds, HRW's single-parameter structure forces it to compromise on both the level and elasticity targets. The additional calibration flexibility of SMCD over HRW greatly improves fit, which is noteworthy given the popularity of HRW among practitioners who seek bounded matching functions (Hagedorn and Manovskii, 2008; Petrosky-Nadeau and Zhang, 2017, 2021).

Table 4: Estimated parameters, pre- and post-October 2008.

	SMCD					CD		HRW	
Period	s	a	p	sa	$\bar{\varepsilon}$	A	ε	ν	$\bar{\varepsilon}$
Pre-2008	0.567	1.762	2.386	1.000	0.465	0.508	0.336	1.154	0.631
Post-2008	0.354	2.829	1.879	1.000	0.311	0.319	0.271	0.622	0.567

Note: All models estimated by NLS in levels. $\bar{\varepsilon}$ = sample mean of $\varepsilon(\theta_t; \hat{\psi})$ over subsample observations (constant for CD).

Table 4 shows how each model absorbs the structural break. For Cobb-Douglas, $\hat{A}$ declines from 0.508 to 0.319 (37%) while $\hat{\varepsilon}$ shifts only from 0.336 to 0.271, confirming the level-shift interpretation from Section 5.1. The SMCD parameters tell the same story: $\hat{s}$ falls 38% (from 0.567 to 0.354), mirroring the decline in $\hat{A}$, while $\hat{a}$ rises and the crossover tightness $1/\hat{a}$ shifts leftward, compressing the vacancy-limited regime post-crisis. The binding constraint $\widehat{sa} = 1.000$ in both periods means NLS allocates the maximum possible vacancy-filling capacity.

The $\bar{\varepsilon}$ column exposes HRW's fundamental limitation most sharply. Pre-2008, the NLS solution lands at $\hat{\nu} = 1.154$, implying a sample-average elasticity of 0.631—roughly double CD's estimate of 0.336. Post-2008, $\hat{\nu}$ collapses to 0.622, and the implied $\bar{\varepsilon} = 0.567$ remains far above CD's 0.271. HRW's ν is not measuring the matching elasticity. It is absorbing the efficiency level at the cost of a severely distorted curvature profile, which explains the large fit gap relative to both SMCD and CD. The SMCD average elasticities (0.465 pre, 0.311 post) exceed CD's constant $\hat{\varepsilon}$ as well, but for a different reason: SMCD's elasticity profile is genuinely nonlinear, rising toward one at low tightness and falling toward zero at high tightness, so recession observations where θ is small pull the arithmetic average above the value that log-linear estimation would assign to the typical observation.

5.3. The Two-Plateau Soft-Min Cobb-Douglas Matching Function

The single-p SMCD dominates HRW conditional on respecting probability bounds, but Cobb-Douglas retains a fit advantage in Table 3 that originates almost entirely from its extrapolation into the $q > 1$ region. A natural one-parameter extension of the SMCD largely eliminates this residual advantage by assigning independent curvature to each matching regime. Define the *two-plateau soft-min Cobb-Douglas* (2P-SMCD) matching function from the job finding rate:

$$f(\theta) = f_c \cdot \frac{\theta}{\theta_c} \left[\frac{2}{1 + (\theta/\theta_c)^p} \right]^{1/p}, \qquad p = \begin{cases} p_L & \theta \leq \theta_c \\ p_R & \theta > \theta_c, \end{cases} \tag{4}$$

with parameters $(f_c,\, \theta_c,\, p_L,\, p_R)$, where $f_c > 0$, $\theta_c > 0$, and $p_L, p_R > 0$. The parameter θ_c is the crossover tightness and $f_c = f(\theta_c)$ is the job-finding rate there. The single-p SMCD is recovered at $p_L = p_R$, with $\theta_c = 1/a$ and $f_c = s/2^{1/p}$. The extension adds exactly one genuinely new parameter: the asymmetry $p_L \neq p_R$. The crossover θ_c was already present in the single-p model as $1/a$.

The two tail caps are independently controlled. As $\theta \to \infty$, $f(\theta) \to s_R \equiv f_c \cdot 2^{1/p_R}$. As $\theta \to 0$, $q(\theta) \equiv f(\theta)/\theta \to q_{\max} \equiv (f_c/\theta_c) \cdot 2^{1/p_L}$. Setting p_L large makes the vacancy-filling probability rise sharply toward $q_{\max}$ in slack markets without imposing any restriction on how quickly the job-finding rate saturates in tight markets, which is governed by p_R alone. This decoupling has a direct economic reading. The parameter p_L governs how quickly vacancy-filling saturates as workers grow scarce, while p_R governs how quickly job-finding saturates as vacancies grow scarce. The crossover θ_c is the balance point at which the two sides of the market contribute equally to matching ($\varepsilon(\theta_c) = 1/2$): to its left, in slack markets, matches respond more to vacancies. To its right, in tight markets, they respond more to searchers. There is no *a priori* reason the pace at which each margin congests should be symmetric across these regimes. The two-plateau relaxes exactly this restriction, at the cost of a single parameter.[9]

[9] A globally smooth curvature profile could accommodate the same asymmetry, but only by adding several parameters and forfeiting both the analytic probability bounds of Proposition 2 and the closed-form inverse q^{-1} that the free-entry step of the DMP exercise relies on (Appendix C). The two-plateau is the most parsimonious departure from constant curvature—a two-segment linear spline in $(\log\theta, \operatorname{logit}\varepsilon)$ space—and, because it strictly nests the single-p form at $p_L = p_R$, the asymmetry is a testable restriction rather than free tuning.

For either branch, $\varepsilon(\theta) = 1/(1 + (\theta/\theta_c)^p)$. Direct computation gives

$$\frac{\varepsilon}{1-\varepsilon} = \frac{1/(1+(\theta/\theta_c)^p)}{(\theta/\theta_c)^p/(1+(\theta/\theta_c)^p)} = \left(\frac{\theta_c}{\theta}\right)^p,$$

and therefore

$$\operatorname{logit}\varepsilon(\theta) \;\equiv\; \log\frac{\varepsilon}{1-\varepsilon} \;=\; -p\log\left(\frac{\theta}{\theta_c}\right).$$

The logit of the matching elasticity is linear in $\log\theta$, with slope $-p$ and intercept zero at θ_c (where $\varepsilon = 1/2$). For the single-p SMCD, a single line with slope $-p$ describes the entire elasticity profile. For the 2P-SMCD, the profile is piecewise linear: slope $-p_L$ for $\theta < \theta_c$ and slope $-p_R$ for $\theta > \theta_c$, joined at $(\log\theta_c, 0)$. The two slopes are separately identified by data on each side of the crossover. Notably, θ_c plays a dual role: it sets the logit-intercepts ($p_{L,R}\log\theta_c$) and marks the kink where curvature shifts between regimes (Figure 4).

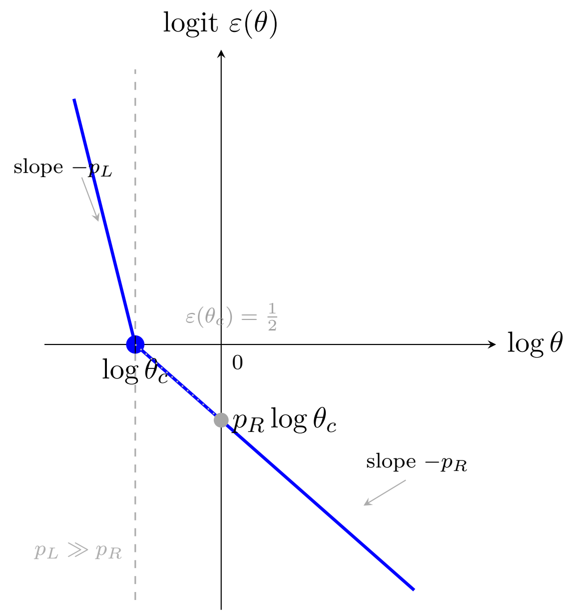

Figure 4: Two-plateau SMCD in $(\log\theta, \text{logit }\varepsilon)$ space (schematic). The log-elasticity is piecewise linear: slope $-p_L$ for $\theta < \theta_c$ and $-p_R$ for $\theta > \theta_c$, joined at $(\log\theta_c, 0)$ where $\varepsilon = 1/2$. Slopes are drawn for legibility with $p_L \gg p_R$. The estimated p_L is far steeper (near-vertical). The dotted line extends the right segment to the y-axis, marking the intercept $p_R\log\theta_c$.

This characterization also clarifies the relationship to the CES discussed in Section 4: the 2P-SMCD produces a piecewise-linear profile, which is outside the CES class entirely. Moreover, the generalization from the single-p form is intuitive: the single-p SMCD forces one curvature parameter to govern both the vacancy-filling cap in slack markets and the job-finding cap in tight markets, while the 2P-SMCD assigns each regime its own curvature.

Proposition 2 confirms that this additional flexibility preserves all four desiderata.

Proposition 2 (Properties of the 2P-SMCD)**.** *The two-plateau matching function defined in* (4) *satisfies the following four properties.*

(i) Constant returns to scale. $m(\lambda u, \lambda v) = \lambda m(u, v)$ *for all* $\lambda > 0$.

(ii) Continuously differentiable and concave. $f(\theta)$ *is* C^1 *on* $(0, \infty)$*, smooth on each branch, and globally concave.*

(iii) Joint calibration. *For any feasible triple* $(f^*, \varepsilon^*, \theta^*)$ *with* $f^* \leq \min\{1, \theta^*\}$ *and* $\varepsilon^* \neq 1/2$*, there exist parameters* (f_c, θ_c, p_R)*—or* (f_c, θ_c, p_L)*—that achieve* $f(\theta^*) = f^*$ *and* $\varepsilon(\theta^*) = \varepsilon^*$*.*

(iv) Bounded probabilities. $f(\theta) \leq s_R$ *and* $q(\theta) \leq q_{\max}$ *for all* $\theta > 0$*. Setting* $s_R \leq 1$ *and* $q_{\max} \leq 1$ *ensures* $f, q \in (0, 1)$*.*

Proof. *(i)* $\theta = v/u$ is homogeneous of degree zero, so $m(\lambda u, \lambda v) = \lambda u \cdot f(\lambda v/\lambda u) = \lambda m(u, v)$. *(ii)* On each open branch, $f''(\theta) < 0$ by the same calculation as Lemma 2: with $z = \theta/\theta_c$, $h''(z) = -(p+1)z^{p-1}(1+z^p)^{-2-1/p} < 0$ for $p = p_L$ on the left and $p = p_R$ on the right. At θ_c, both branches yield $f(\theta_c) = f_c$. Differentiating and evaluating at $z = 1$ gives $f'(\theta_c) = (f_c/\theta_c) \cdot 2^{1/p} \cdot 2^{-1/p-1} = f_c/(2\theta_c)$, which is independent of p. Hence $f'(\theta_c^-) = f'(\theta_c^+)$ and f is C^1. A C^1 function that is strictly concave on each of two abutting intervals is globally concave (the secant inequality decomposes at the junction into two within-branch inequalities). *(iii)* For $\varepsilon^* < 1/2$ (calibration point on the right branch, $\theta^* > \theta_c$): the elasticity condition gives $(\theta^*/\theta_c)^{p_R} = (1-\varepsilon^*)/\varepsilon^*$, which pins p_R given θ_c. The level condition then gives $f_c = f^*/(2(1-\varepsilon^*))^{1/p_R}$. The case $\varepsilon^* > 1/2$ is symmetric on the left branch. *(iv)* On the right branch, $f(\theta) = f_c(\theta/\theta_c)[2/(1+(\theta/\theta_c)^{p_R})]^{1/p_R} \leq f_c \cdot 2^{1/p_R} = s_R$. The left branch and the vacancy-filling bound are analogous. □

Local Cobb-Douglas behavior follows directly from (ii) and (iii). Since $\varepsilon(\theta)$ is continuous everywhere, including at θ_c, then at any calibration point θ^* where $\varepsilon(\theta^*) = \varepsilon^*$, this yields $d\log m = (1-\varepsilon^*)d\log u + \varepsilon^* \, d\log v$ to first order. Thus, the 2P-SMCD is locally Cobb-Douglas at θ^*. The second-order deviation is $-\frac{1}{2}p_{\text{branch}}\, \varepsilon^*(1-\varepsilon^*)\Delta^2$, where p_{branch} is the curvature of the branch containing θ^*.

In the single-p SMCD, a single parameter p must satisfy both bound constraints simultaneously. At the US calibration the vacancy-filling bound forced $p_{\min} \approx 2.21$, producing curvature $|\kappa(\theta^*)| \approx 0.46$ even though the calibration point lies in the worker-limited regime. The 2P-SMCD decouples the two constraints: the right cap $s_R \leq 1$ is governed by p_R alone, while p_L independently controls the left cap $q_{\max} \leq 1$. The minimum feasible p_R—the smallest value for which a finite p_L satisfying $q_{\max} \leq 1$ exists—is approximately 0.97, reducing curvature at θ^* to $|\kappa| \approx 0.20$, roughly half the single-p value. The corresponding right-branch ceiling is $s_R \approx 0.569$, a 22% improvement over the single-p cap of $s \approx 0.467$. Figure 5 plots all four specifications at the common calibration. In the empirical horse race of Section 5.4, the NLS estimates are unconstrained by the sample-mean calibration, so the estimated s_R is free to exceed this theoretical ceiling.

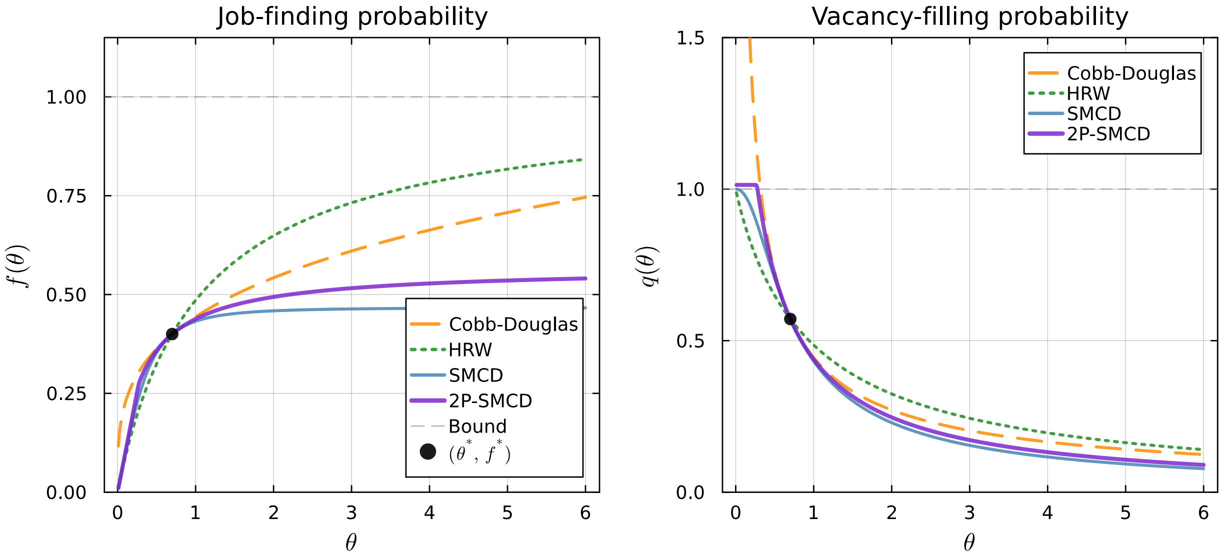

Figure 5: Job-finding probability $f(\theta)$ (left) and vacancy-filling probability $q(\theta)$ (right) for CD (dashed), HRW (dotted), single-p SMCD (solid, light), and 2P-SMCD (solid, dark). All calibrated to $(\theta^*, f^*, \varepsilon^*) = (0.70, 0.40, 0.29)$ where feasible. HRW matches f^* at θ^* only. Filled circle: calibration point.

5.4. Horse Race: Adding the Two-Plateau SMCD

The 2P-SMCD is estimated by nonlinear least squares. We impose the probability bounds $s_R \equiv f_c 2^{1/p_R} \leq 1$ and $q_{\max} \equiv (f_c/\theta_c)2^{1/p_L} \leq 1$ by parametrizing f_c as a fraction of its cap, so the $q > 1$ violation rate is zero by construction. The left-branch curvature p_L is identified only from below: once the branch is steep enough for near-Leontief behavior, further steepness leaves the fit essentially unchanged. Rather than estimate p_L freely—which drives it to the hard-min limit—we calibrate it to the smoothest value within 0.5% of the best attainable fit and estimate (θ_c, f_c, p_R) conditional on it (Appendix D.1). As in the single-p horse race, Table 5 reports a restricted RMSE* alongside the full-sample RMSE, computed on the admissible domain $\theta > \theta_{\min}$ of Section 5.2 ($\theta_{\min} = 0.36$ pre-break, 0.21 post-break), where CD also predicts valid probabilities. Table 6 reports the estimates and derived caps.

Table 5: Fit statistics for the 2P-SMCD and Cobb-Douglas, pre- and post-October 2008.

Period	Model	Par	RMSE*	RMSE	MAE	q>1 (%)
Pre-2008	CD	2	0.0447	0.0442	0.0349	12.0
	2P-SMCD	4	0.0457	0.0499	0.0397	0.0
Post-2008	CD	2	0.0293	0.0284	0.0224	9.4
	2P-SMCD	4	0.0283	0.0285	0.0220	0.0

Note: $n = 693$ pre-break, $n = 203$ post-break. RMSE* computed on months with $\theta > \theta_{\min}$ ($\theta_{\min} = 0.36$ pre-break, 0.21 post-break), where CD also satisfies $q \leq 1$. CD RMSE* exceeds its full-sample RMSE because the excluded low-θ months are recession observations where CD's unconstrained extrapolation fits well.

Table 6: Estimated 2P-SMCD parameters, derived tail caps, and average tightness elasticity, pre- and post-October 2008.

Period	θ_c	f_c	p_L	p_R	s_R	$q_{\max}$	$\bar{\varepsilon}$
Pre-2008	0.355	0.336	12.5	0.675	0.938	1.000	0.439
Post-2008	0.202	0.189	10.5	0.864	0.422	1.000	0.309

Note: $s_R = f_c \cdot 2^{1/p_R}$; $q_{\max} = (f_c/\theta_c) \cdot 2^{1/p_L}$. $\bar{\varepsilon} = T^{-1} \sum_t \varepsilon(\theta_t; \hat{\psi})$: average tightness elasticity over the subsample, from the piecewise analytical formula (Section 5.3).

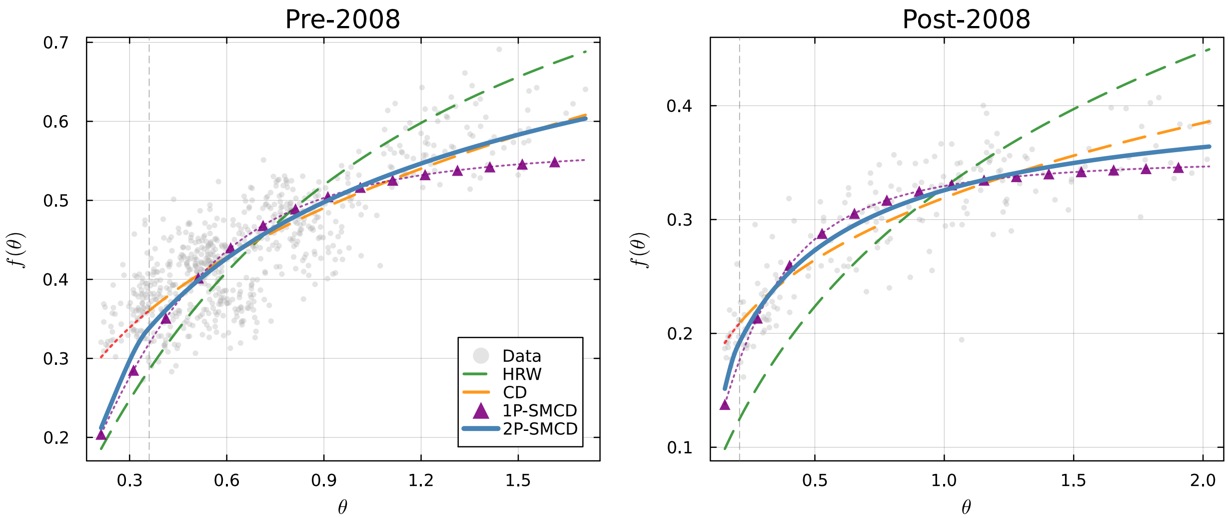

Figure 6: Job-finding rate $f(\theta)$ for all four specifications—2P-SMCD (bold solid blue), 1P-SMCD (purple triangles), HRW (dashed green), and CD (dashed orange)—pre-break (left) and post-break (right). Grey scatter: data. Red dotted segment: CD in the region $\theta < \theta_{\min}$ where fitted $q > 1$. Vertical gray line: $\theta_{\min}$. Legend in the left panel.

Tables 5–6 and Figure 6, which overlays all four specifications, deliver three findings.

First, the two plateaus strictly improve on the single-p SMCD in both subsamples—RMSE 0.0499 versus 0.0564 pre-break and 0.0285 versus 0.0313 post-break—and close most of the remaining distance to Cobb-Douglas. On the full time series CD still posts a marginally lower RMSE, because it is free to extrapolate into the $q > 1$ region, but post-break the two are effectively tied (0.0284 versus 0.0285).

Second, as in the single-p horse race (Section 5.2), CD's edge is concentrated in months where it is not a valid probability model. The full-sample RMSE gap (2P-SMCD minus CD) is $0.0499 - 0.0442 = 0.0057$ pre-break and a negligible $0.0285 - 0.0284 = 0.0001$ post-break. On the admissible domain $\theta > \theta_{\min}$, where CD also respects $q \leq 1$, the gap nearly closes pre-break (RMSE* of 0.0457 versus 0.0447, a 2.2% difference) and *reverses* post-break, where the 2P-SMCD fits better than CD (0.0283 versus 0.0293). CD's advantage thus lives entirely in the recession months where it predicts a vacancy-filling probability above one.

Third, the estimated shape is strongly asymmetric: p_L is an order of magnitude larger than p_R (12.5 versus 0.675 pre-break, 10.5 versus 0.864 post-break). The vacancy-filling rate rises sharply to its cap on the left branch while the job-finding rate saturates gradually on the right, a shape a single-p specification cannot produce. The vacancy-filling bound binds at the optimum in both subsamples ($q_{\max} = 1$), which is precisely why unconstrained CD overshoots $q > 1$ in recessions. As described above, p_L is identified only from below and is calibrated rather than estimated (Appendix D.1). Conditional on this value, θ_c and p_R are sharply identified (95% profile intervals $\theta_c \in [0.34, 0.36]$, $p_R \in [0.65, 0.90]$ pre-break and

$[0.17, 0.22]$, $[0.64, 1.01]$ post-break), while f_c sits on the binding vacancy-filling cap.

The structural break carries a direct message for applied users. Researchers should use the sub-period estimates in Tables 5 and 6 corresponding to their application's temporal focus. Full-sample estimates are defensible only if the structural model endogenously explains the efficiency decline through search intensity, worker composition, or mismatch (Şahin et al., 2014; Barnichon and Figura, 2015). A further simplification is that the vacancy-filling bound binds at the estimated optimum in every split we examine: $q_{\max} \equiv (f_c/\theta_c)\, 2^{1/p_L} = 1$ (Table 6). The junction level is then no longer free: $f_c = \theta_c\, 2^{-1/p_L}$. Substituting into (4), the 2P-SMCD collapses to a three-parameter specification in (θ_c, p_L, p_R),

$$f(\theta) = 2^{-1/p_L}\, \theta \left[\frac{2}{1 + (\theta/\theta_c)^p}\right]^{1/p}, \qquad p = \begin{cases} p_L, & \theta \leq \theta_c, \\ p_R, & \theta \geq \theta_c. \end{cases} \tag{5}$$

The bound binds because U.S. recessions push tightness low enough that the fit reaches the vacancy-filling ceiling. Researchers modeling the U.S. labor market can therefore adopt (5) directly, plugging in the values of (θ_c, p_L, p_R) from Table 6 for the sub-period matching their application.

6. A Quantitative DMP Exercise

The horse race establishes that the 2P-SMCD fits matching data as well as CD while respecting probability bounds, and strictly dominates HRW. Here we ask how much the distinction between the two bounded specifications matters for efficiency and unemployment crises. We embed the pre-break 2P-SMCD estimates in a DMP economy with Pissarides (2009) fixed matching costs and compare its business-cycle behavior to an otherwise identical HRW economy. Both are calibrated by simulated method of moments to the *same* three targets—the relative volatility of unemployment, mean tightness, and the wage elasticity—so any remaining difference is attributable solely to the *shape* of the matching function.

Three results emerge. First, bounded matching probabilities imply a modified Hosios condition (Proposition 3) under which the efficient bargaining weight $\alpha_L^H(\theta)$ depends on tightness and on the fixed-cost share ϕ, collapsing toward zero as $\phi \to 1$. This is a new analytical result: under the standard (flow-cost-only) Hosios condition, the efficient weight is the constant $\eta_u(\theta)$. With the two-part cost of Pissarides (2009), it becomes $\alpha_L^H(\theta) = (1-\phi(\theta))\eta_u(\theta)/[1-\phi(\theta)\eta_u(\theta)]$. A planner allocation reveals a sign-reversing unemployment gap: the 2P-SMCD economy is inefficiently tight on average but generates 11.1 pp of excess unemployment in crises, while HRW sits near constrained efficiency. Second, the near-Leontief recession branch generates vacancy shutdowns and fat-tailed crisis episodes—the

Petrosky-Nadeau and Zhang (2017) channel, which persists under a matching function *jointly calibrated* to level and elasticity targets. Third, the Hosios wedge is roughly twice as volatile under the 2P-SMCD as under HRW, because its steeper elasticity gradient amplifies the gap between the constant bargaining weight and the state-dependent efficient share.

This exercise requires a global solution for the following reasons. First, the job-finding rate is bounded and concave, so unemployment rises faster in recessions than it recedes in expansions. A first-order approximation is symmetric by construction, so it misses this asymmetry. More subtly, higher-order perturbation does not repair this problem, because the equilibrium tightness policy is not analytic. In particular, the vacancy non-negativity constraint makes firms stop posting once a filled match is worth less than the hiring cost. This puts a kink in the tightness policy at the shutdown threshold. The two-plateau crossover θ_c leaves the policy only once-differentiable at a second point. Both points lie inside the ergodic set the economy visits. A Taylor expansion around the deterministic steady state therefore has a radius of convergence bounded by the distance to these singularities. Adding orders can actually hinder accuracy in these circumstances, as Petrosky-Nadeau and Zhang (2017).

We therefore solve the model using time iteration on the job-creation condition. We use the Rouwenhorst method to construct the productivity grid. To better capture the tail statistics, the continuation value is then evaluated by Gauss–Hermite quadrature over the continuous productivity distribution, rather than the associated discrete transition matrix. Following Petrosky-Nadeau and Zhang (2021), we estimate the structural parameters on the globally simulated ergodic distribution, so the same nonlinear dynamics that generate the results also discipline the calibration.

6.1. Environment and equilibrium

Time is discrete and a period is one month. A filled job produces output z_t, where $\log z_t$ follows an AR(1) process.[10] Unemployed workers and vacancies meet according to the matching function $m(u, v)$, generating a job-finding rate $f(\theta) = m(1, \theta)$ and a vacancy-filling rate $q(\theta) = m(1, \theta)/\theta$ at tightness $\theta = v/u$. Matches dissolve at the exogenous monthly rate δ.

Firms post vacancies subject to the two-part hiring cost of Pissarides (2009): a flow cost k for each period a vacancy is open, plus a fixed cost κ paid once a match forms.[11] The expected cost of a hire is $k/q(\theta) + \kappa$, and we summarize its composition by the fixed-cost

[10] We estimate the productivity process on the pre-break sample: quarterly HP-filtered log output per worker has autocorrelation 0.757 and innovation standard deviation 0.0084, which map to monthly $\rho_z = 0.911$ and $\sigma_z = 0.0053$. The conditional expectation is evaluated by Gauss–Hermite quadrature over the continuous innovation distribution. See Section 6.3 and Appendix C.

[11] Here κ denotes the fixed matching cost, distinct from the log-curvature $\kappa(\theta)$ of Section 4.

share

$$\phi(\theta) = \frac{\kappa}{\kappa + k/q(\theta)}. \tag{6}$$

Free entry drives the value of an open vacancy to zero whenever vacancies are posted, $q(\theta_t)\left(J_t - \kappa\right) = k$: a firm pays the flow cost k to hold the vacancy open and, upon meeting a worker at rate $q(\theta_t)$, pays the fixed cost κ and obtains a filled job worth J_t. The fixed cost is incurred immediately upon matching, so it enters undiscounted. This timing feature accords with Pissarides (2009) and Christiano et al. (2016), whose free-entry condition is likewise $q\,(J - \kappa)$ equals the posting cost.[12] The value of a filled job satisfies the Bellman equation

$$J_t = \beta\,\mathbb{E}_t[z_{t+1} - w_{t+1} + (1-\delta)\,J_{t+1}]\,, \tag{7}$$

where β is the monthly discount factor. The value of an open vacancy is $V_t^v = -k + q(\theta_t)\big(J_t - \kappa\big)$. Free entry drives it to at most zero, $V_t^v \leq 0$, or equivalently

$$J_t \leq \kappa + \frac{k}{q(\theta_t)}, \qquad \theta_t \geq 0, \qquad \theta_t\left(\kappa + \tfrac{k}{q(\theta_t)} - J_t\right) = 0. \tag{8}$$

Whenever tightness is positive, free entry holds with equality, $J_t = \kappa + k/q(\theta_t)$, and combining it with (7) yields the job-creation condition

$$\kappa + \frac{k}{q(\theta_t)} = \beta\,\mathbb{E}_t\left[z_{t+1} - w_{t+1} + (1-\delta)\left(\kappa + \frac{k}{q(\theta_{t+1})}\right)\right], \tag{9}$$

which equates the hiring cost of a filled job (left-hand side) to its present discounted value (right-hand side). Because $q(\theta) \leq q_{\max} < \infty$, the cost floor $\kappa + k/q_{\max}$ cannot be undercut: once J_t falls below it no firm posts and tightness hits the corner $\theta_t = 0$—the second and third parts of (8), a treatment of the vacancy non-negativity constraint in the manner of Petrosky-Nadeau and Zhang (2017).

Wages are set by Nash bargaining with worker weight α_L,

$$w_t = (1-\alpha_L)\,b + \alpha_L\left(z_t + k\,\theta_t\right), \tag{10}$$

[12] Our monthly timing separates matching from production: a match formed in t first produces in $t+1$, so J_t in (7) is a forward, discounted value. Christiano et al. (2016) let matching and production coincide within the quarter. In both cases κ is paid undiscounted at the match.

with b the flow value of nonemployment. Unemployment evolves as

$$u_{t+1} = u_t + \delta(1 - u_t) - f(\theta_t)\, u_t. \tag{11}$$

Given the exogenous productivity process $\{z_t\}$, an equilibrium is a set of sequences $\{\theta_t, u_t, w_t\}$, with $\theta_t \geq 0$, satisfying the free-entry complementarity (8), the wage equation (10), and the law of motion (11) in every period. As productivity is the only state that bears on decisions, tightness and the wage can be represented as time-invariant functions $\theta_t = \theta(z_t)$ and $w_t = w(z_t)$. We solve for these policy functions globally (Section 6.3).

This shutdown corner is absent from the unbounded Cobb-Douglas benchmark, where $q(\theta) \to \infty$ as $\theta \to 0$ and the cost floor collapses to κ. Under a bounded q—whether 2P-SMCD or HRW—it is a live event, where a capped filling rate leaves a deep recession no way to fill jobs faster. Yet the fat right tail of unemployment in Section 6.4 does not hinge on the corner itself. In the near-Leontief regime around it, the flat q makes tightness and job finding fall convexly with productivity, and the depressed job-finding rate causes unemployment to rise. The shutdown is the limiting case of this contemporaneous nonlinearity, not a separate anticipatory force.

6.2. Efficiency: the Hosios condition with fixed matching costs

Section 4.1 established that bounded probabilities make the Hosios condition $\alpha_L = \eta_u(\theta)$ state-dependent: it can hold at most one point of the cycle, and the resulting wedge is countercyclical. We now generalize this result to account for the fixed matching cost κ of the present environment, which modifies the efficient bargaining weight. The quantitative analysis then shows that this produces a sign-reversing unemployment gap.

Proposition 3. *The decentralized equilibrium is constrained efficient if and only if the worker's bargaining weight satisfies*

$$\alpha_L \;=\; \alpha_L^H(\theta) \;\equiv\; \frac{\big(1 - \phi(\theta)\big)\, \eta_u(\theta)}{1 - \phi(\theta)\, \eta_u(\theta)}, \tag{12}$$

with $\phi(\theta)$ given by (6). At $\phi = 0$ this reduces to the standard condition $\alpha_L = \eta_u(\theta) = 1 - \varepsilon(\theta)$ of Section 4.1.

The intuition is that only the flow component $k/q(\theta)$ of the hiring cost is congestion-sensitive: it is the margin through which an extra vacancy lengthens the queue and lowers q for other firms. The elasticity of the total hiring cost $k/q(\theta) + \kappa$ with respect to tightness is $(1 - \phi(\theta))\, \eta_u(\theta)$—the flow-cost congestion elasticity η_u diluted by the flow share $1 - \phi$. The fixed matching cost κ is paid per match irrespective of market congestion and therefore carries no externality. As ($\phi \to 1$), the congestion externality vanishes and the efficient worker share falls to zero. The derivation of the dynamic social planner's problem with the two-part cost is in Appendix B.2. The standard condition $\alpha_L = \eta_u(\theta)$ obtains in the flow-cost-only environment of Pissarides (2000, ch. 8). To the best of our knowledge, (12) has not previously been stated. Mangin and Julien (2021) generalize the Hosios condition along a different margin—nonlinear surplus splitting—while retaining a single flow posting cost. Their modification and ours are complementary and could in principle be combined.

6.3. Calibration and solution

We adopt the estimates for the matching-function parameters from Section 5.4 at the pre-break period: the 2P-SMCD parameters (θ_c, f_c, p_L, p_R) and the HRW parameter ν (with $s = a = 1$). This cleanly separates identification: the shape of the matching technology is disciplined by the matching data, while the three structural parameters (b, α_L, k) are estimated inside the model. The separation rate is $\delta = 0.035$ (Shimer, 2005), the monthly discount factor is $\beta = e^{-0.04/12}$. We set the fixed-cost share to $\phi = 0.90$ at the steady state, pinning κ from k and $q(\theta^*)$ via (6). A share this high is motivated by Christiano et al. (2016), who estimate a search-theoretic DSGE model by Bayesian impulse response matching and find the fixed component accounts for 94% of the cost of meeting a worker. Our 90% is a conservative lower bound of their estimate.

We use three moments from the simulated data to identify the parameter triple (b, α_L, k): the relative volatility of unemployment $\sigma(u)/\sigma(z) = 10.8$, the mean tightness $\bar{\theta} = 0.67$, and the wage elasticity $\varepsilon_{w,y} = 0.52$. The flow value b governs the size of the match surplus and hence the amplification of unemployment to productivity through the mechanism of Hagedorn and Manovskii (2008). The flow cost k sets the level of tightness—and thereby the average operating point on the matching function. The worker's bargaining weight α_L controls how strongly wages track productivity.

Volatility is measured as HP-filtered *proportional* deviations from the mean, following Petrosky-Nadeau and Zhang (2017), rather than log deviations. Although unemployment is bounded away from zero in the data and in the model's ergodic distribution—so the log transform is well-defined for u—the proportional-deviation convention preserves the cardinal interpretation of rates and maintains a single consistent metric across all variables, including vacancies and tightness, which do reach zero at the shutdown corner.[13] By targeting $\bar{\theta}$ and

[13]The two transforms agree to first order and diverge only in the tails: the log stretches the left tail ($\log(x/\bar{x}) \to -\infty$ as $x \to 0$) and compresses the right. It therefore inflates $\sigma(\log\theta)$, since tightness collapses toward zero in crises, and deflates $\sigma(\log u)$, since unemployment's right tail fattens, biasing $\sigma(\log\theta)/\sigma(\log u)$ upward. Proportional deviations impose no such curvature, so the measured asymmetry is the model's, not the metric's.

$\sigma(u)/\sigma(z)$ rather than mean unemployment and $\sigma(\theta)/\sigma(z)$, we let both the mean unemployment rate and the tightness volatility ratio be *outcomes* that can differ across specifications, revealing how the matching function's shape affects both the level and the higher moments of unemployment.

Targeting $\bar{\theta}$ rather than $\bar{u}$ implies that the model will overshoot the data's mean unemployment. Given the horse-race estimates, $f(0.67) \approx 0.44$–0.45. Flow balance pins steady-state unemployment at $u^* = \delta/(\delta + f(\bar{\theta})) \approx 7.3\%$—already above the data's 5.6%. The ergodic mean lies weakly above even this steady-state value: because $f(\theta)$ is concave, Jensen's inequality implies $\mathbb{E}[f(\theta)] < f(\mathbb{E}[\theta])$, so workers find jobs less efficiently on average than they would at the average level of tightness, pushing $\bar{u}$ above u^*—consistent with the simulated 7.4–7.6%.

The gap between the model and empirical mean of unemployment reflects the two-state model's coarseness: all transitions flow through a single separation rate $\delta = 0.035$ and a single job-finding rate f, abstracting from job-to-job moves and participation-margin flows that lower observed unemployment. Any DMP model with these primitives produces $\bar{u} \approx 7$–7.5%. If we instead targeted $\bar{u} = 5.6\%$, the model would need $\bar{f} \approx 0.59$, above the average rate implied by the horse-race estimates. The overshoot is the price of letting the matching data discipline the mean rather than the unemployment rate.

The amplification target $\sigma(u)/\sigma(z) \approx 10.8$ follows the convention of Petrosky-Nadeau and Zhang (2017) for quantifying the unemployment-volatility puzzle of Shimer (2005). In a single-shock model the only question is which unconditional volatility ratio the shock should reproduce. Matching a conditional elasticity—attributing to the model only the productivity-driven share of fluctuations—would deliver a far tamer economy in which the global nonlinearities that separate the 2P-SMCD from HRW would scarcely operate. We reach the amplification regime not through a near-productivity value of nonemployment, as in Hagedorn and Manovskii (2008), but through the fixed matching cost of Pissarides (2009): the estimated flow value of nonemployment, $b \approx 0.81$ for the 2P-SMCD (Table 7), lies well below the Hagedorn–Manovskii value of 0.955—which Hall and Milgrom (2008) show implies a Frisch labor-supply elasticity outside any empirical range—and near the 0.71 that Hall and Milgrom obtain by imputing the value of non-work directly.[14] The amplification is the fixed matching cost's doing, not a thin surplus wrung from a near-productivity outside option.

[14]These figures are flow values of nonemployment relative to labor productivity ($\bar{z} = 1$), the object comparable across studies. They are *not* replacement ratios. The implied replacement ratio b/w is higher still, because the low estimated bargaining weight ($\alpha_L \approx 0.11$) leaves the wage only slightly above the outside option.

Solution algorithm. The model is solved by time iteration on the job-creation condition using Gauss–Hermite quadrature over the continuous log-normal innovation distribution, eliminating the discretization staircase that a finite Markov chain introduces in the tails. Given a guess for the expected continuation value $E(z) = \mathbb{E}[J(z') \mid z]$, each iteration inverts the free-entry condition (9) for $\theta(z)$. If the implied value falls short of $\kappa + k/\bar{q}$ the vacancy-shutdown corner binds and $\theta(z) = 0$.

The three parameters are estimated by simulated method of moments (SMM) on the global solution.[15] We solve for the bargaining weight α_L apart from the two cost parameters, because each target has a parameter it speaks to most directly: b governs the amplification of unemployment, through the size of the match surplus. k governs the level of tightness, through the level of vacancy posting. α_L governs the cyclicality of wages, through the share of the procyclical surplus that workers capture. We put this structure to work by holding α_L fixed, solving for the cost pair (b, k) that reproduces $\sigma(u)/\sigma(z)$ and $\bar{\theta}$, and then turning the single dial α_L—a higher weight makes wages more procyclical—until the simulated wage elasticity matches the data, re-solving (b, k) at each turn. Appendix C details the solver and the estimation loop. The calibration and the reported second moments are computed on a 200,000-month simulation. The high-variance tail statistics—level skewness, crisis frequency, and the crisis gap—are instead reported from a 5-million-month simulation, at which they converge and the seed is immaterial (Appendix D.4). Table 7 reports the estimated parameters and their fit to the three targets.

6.4. Results

Table 8 reports business-cycle moments for the two calibrated economies alongside the pre-break data. By construction both match $\sigma(u)/\sigma(z)$, $\bar{\theta}$, and the wage elasticity. The differences lie in untargeted objects. Mean unemployment, now an outcome, is 7.4% under the 2P-SMCD and 7.6% under HRW—both above the data's 5.6%, for the reasons discussed in Section 6.3. More revealing are the untargeted second moments. Both economies produce a tightness volatility ratio near 17: ($\sigma(\theta)/\sigma(z) = 17.4$ under the 2P-SMCD, 17.1 under HRW), below the data value of 21.4 but with the 2P-SMCD marginally closer. The two specifications separate on the job-finding rate, which is markedly less volatile under the 2P-SMCD ($\sigma(f)/\sigma(z) = 8.2$ versus 11.0), because its bounded upper plateau saturates f in booms while HRW's smoother profile keeps the matching elasticity relatively high throughout. The unemployment–productivity correlation is weaker under the 2P-SMCD (-0.58

[15] The three-moment system is exactly identified, so the estimator matches the targets exactly. Following the quantitative DMP literature (Hagedorn and Manovskii, 2008; Petrosky-Nadeau and Zhang, 2017), we treat the three data moments as calibration targets and do not conduct inference.

Table 7: DMP calibration: pre-break targets and calibrated parameters.

	2P-SMCD	HRW	Target
Flow value of nonemployment b	0.813	0.698	—
Worker bargaining weight α_L	0.112	0.061	—
Flow vacancy cost k	0.256	0.454	—
Fixed matching cost κ	3.46	6.23	—
Fixed-cost share ϕ	0.90	0.90	0.90
$\sigma(u)/\sigma(z)$	10.83	10.85	10.77
Mean tightness $\bar{\theta}$	0.672	0.670	0.670
$\varepsilon_{w,y}$	0.53	0.52	0.52

Note: Monthly model. $\delta = 0.035$, $\beta = e^{-0.04/12}$. Moments in HP-filtered proportional deviations from the mean. All three moments are targeted. $\sigma(\theta)/\sigma(z)$ and $\bar{u}$ are untargeted outcomes (Table 8).

versus -0.80): its state-dependent, asymmetric u dynamics are poorly summarized by a linear comovement, echoing the weaker correlations Petrosky-Nadeau and Zhang (2017) obtain from a global solution relative to log linearization and closer to the mild relationship in the data.

Two caveats apply to the comparison with data. First, both economies are driven by a single productivity shock, so model variables comove nearly perfectly with z, whereas the data, presumably driven by many shocks, show weak correlations with productivity. This also explains why $\sigma(w)/\sigma(z) = 0.53$ under the 2P-SMCD and 0.52 under HRW fall below the empirical 0.93 despite matching the wage elasticity $\varepsilon_{w,y}$ in each case, since $\rho(w, z) \approx 1$ makes relative volatility and elasticity coincide in the model. Second, as the single shock must account for all unemployment variation, the calibrated b is higher than micro estimates would suggest. A multi-shock model would spread the amplification burden across sources.

The unemployment distribution. The sharpest difference is in the shape of the unemployment distribution. Figure 7 plots the simulated density of the monthly unemployment rate under the two calibrations. Panel (a) shows that the 2P-SMCD core is *more concentrated* than HRW's: its unconditional standard deviation is 0.93 percentage points against 1.05, and its mode is taller and narrower. HRW carries visibly more mass in the range 9–12%. A reader who stopped at panel (a) would conclude that HRW generates more unemployment risk. Panel (b), plotted on a log scale, corrects that impression. Beyond about 12% the 2P-SMCD density dominates, and by the 15% crisis threshold of Petrosky-Nadeau and Zhang (2021) the 2P-SMCD density is roughly ten times higher (the crisis *frequency* is 0.22%

Table 8: Business-cycle moments: data and calibrated models.

		Data	2P-SMCD	HRW
Tightness θ	$\sigma(x)/\sigma(z)$	21.36	17.39	17.09
	$\rho(x,z)$	0.30	1.00	1.00
	autocorr	0.88	0.70	0.70
Unemployment u	$\sigma(x)/\sigma(z)$	10.77	10.83	10.85
	$\rho(x,z)$	−0.29	−0.58	−0.80
	autocorr	0.88	0.74	0.78
	mean (level)	5.62%	7.42%	7.59%
Vacancies v	$\sigma(x)/\sigma(z)$	11.44	12.98	10.60
	$\rho(x,z)$	0.40	0.93	0.88
	autocorr	0.91	0.53	0.42
Job finding f	$\sigma(x)/\sigma(z)$	6.57	8.23	11.02
	$\rho(x,z)$	0.28	0.91	0.99
	autocorr	0.79	0.65	0.70
Wage w	$\sigma(x)/\sigma(z)$	0.93	0.53	0.52
	$\rho(x,z)$	0.62	1.00	1.00
	autocorr	0.80	0.70	0.70

Note: HP-filtered proportional deviations from the mean ($\lambda = 1600$). Data: pre-break subsample, reference variable output per worker. z is model productivity. Model: 200,000-month simulations for the second moments; 5,000,000-month for the tail statistics (skewness, crisis frequency, crisis gap), which converge only at length (Appendix D.4). Mean unemployment (level) is an untargeted outcome.

versus 0.05%, a factor of 4.1). The skewness of the unemployment rate in levels—10.7 for the 2P-SMCD against 1.8 for HRW (Figure 7)—summarizes this reversal: the 2P-SMCD distribution is tight in the core but carries a long, heavy right tail. The two economies have nearly identical mean unemployment (7.4% and 7.6%), so the fixed crisis threshold sits at essentially the same distance above each mean. Skewness, moreover, normalizes by each distribution's own mean and standard deviation, isolating the shape of the right tail from the level and from the choice of threshold.

The mechanism is the near-Leontief left branch. In normal times ($\theta > \theta_c$), the bounded vacancy-filling rate saturates and hiring operates near capacity, compressing the core distribution. But when a sufficiently adverse shock pushes θ below θ_c, the flat q leaves no room for faster hiring, and unemployment spills upward along a near-deterministic path: during vacancy shutdown ($\theta = 0$), the law of motion $u_{t+1} = u_t + \delta(1 - u_t)$ drives unemployment up by roughly $\delta \approx 3.5$ percentage points per month. This produces the secondary peaks visible in panel (b) at approximately 3-percentage-point intervals—the "stepping stones" of one-, two-, and three-month shutdown spells. HRW's smoother matching technology, by contrast, spreads volatility evenly across the cycle: moderate shocks produce moderate unemployment fluctuations throughout, generating broader everyday variation but no catastrophic collapses. Its density decays monotonically with no secondary structure. This is the crisis mechanism of Petrosky-Nadeau and Zhang (2017) and Petrosky-Nadeau and Zhang (2021), arising from the geometry of the matching function rather than from an extreme surplus calibration.

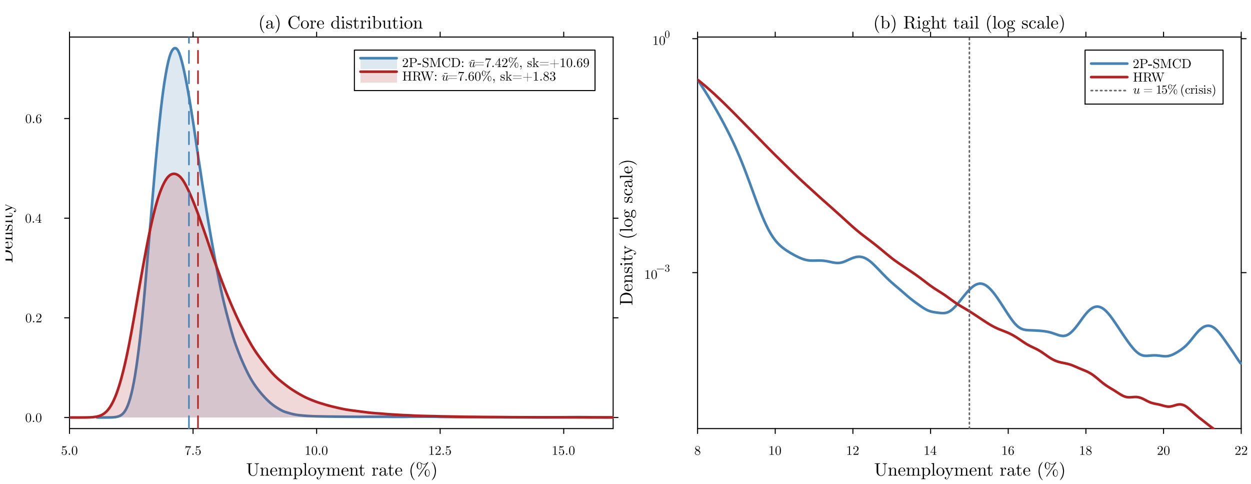

Figure 7: Simulated density of the monthly unemployment rate: 2P-SMCD and HRW. Panel (a): linear scale. Dashed lines mark the means. Panel (b): log scale. Dotted line marks the 15% crisis threshold.

The Hosios wedge. At the calibration, fixed costs dominate hiring ($\phi = 0.90$), so the efficient worker share of Proposition 3 is low—but the two matching functions deliver very different levels. The simulated mean $\alpha_L^H(\theta)$ is 0.134 under the 2P-SMCD and 0.060 under HRW, close to but not identical to the calibrated weights $\alpha_L = 0.112$ and 0.061. On average the decentralized SMCD economy modestly *under*-rewards workers (mean modified wedge $\alpha_L - \alpha_L^H = -0.021$) while HRW sits essentially at the modified Hosios boundary (wedge +0.001), because its elasticity $\eta_u(\theta)$ varies too little over the cycle to open a sizeable wedge. This near-efficiency verdict requires accounting for cost composition: measured against the conventional flow-cost rule $\alpha_L = \eta_u(\theta)$, whose simulated mean is ≈ 0.60 (2P-SMCD) and 0.38 (HRW), the same calibrated $\alpha_L \approx 0.1$ appears to grossly *under*-reward workers—a mean wedge near -0.5—implying wildly excessive vacancy creation. The fixed-cost correction of Proposition 3, recognizing that the 90%-dominant matching cost carries no congestion externality, lowers the efficient share to ≈ 0.1 and overturns that verdict.

The two specifications differ sharply in the *cyclicality* of the wedge. The efficient share

$\alpha_L^H(\theta)$ moves over the cycle while the bargained α_L is fixed. The standard deviation of $\alpha_L - \alpha_L^H(\theta)$ is 0.032 under the 2P-SMCD against 0.018 under HRW. The 2P-SMCD compounds the state-dependence of $\eta_u(\theta)$ with that of $\phi(\theta)$, roughly doubling the cyclical variation. Because $\eta_u(\theta)$ falls toward zero as tightness collapses, the efficient share drops below the bargained α_L in deep recessions. The fraction of time in which the wedge flips sign—workers *over*-appropriate—is 25.2% under the 2P-SMCD against 54.2% under HRW, but HRW's crossings are economically inconsequential (it hovers near zero wedge), while the SMCD's sign flips are concentrated in deep recessions, where over-appropriation coincides with already-high unemployment.

The unemployment gap. The Hosios wedge in parameter space (α_L versus α_L^H) does not by itself quantify the degree of excess or insufficient unemployment. To answer that question, we solve the planner's allocation. This consists of swapping α_L in the model with the (state-dependent) $\alpha_L^H(\theta)$ of Proposition 3, fixing the parameters at the calibrated values, and solving the resulting planner allocation as a full dynamic global solution using the same time-iteration solver, Gauss–Hermite quadrature, and shock realizations as the decentralized economy. The per-period gap $u_t^{\text{dec}} - u_t^{\text{pln}}$ then isolates the welfare cost of constant-α_L bargaining.

Under the 2P-SMCD, the unemployment gap flips sign. On average, the decentralized economy under-bargains ($\alpha_L = 0.112 < \alpha_L^H \approx 0.134$), so firms over-hire and unemployment is *too low*: mean $\bar{u}^{\text{dec}} = 7.4\%$ against $\bar{u}^{\text{pln}} = 7.6\%$, a gap of -0.19 percentage points. But in crisis episodes ($u \geq 15\%$), the mean gap flips to $+11.1$ percentage points of *excess* unemployment—larger than the 9.2-pp peak gap in the representative (median-peak) episode below, because the distribution of crisis peaks is right-skewed and a few extreme spells pull the mean above the median peak. In the near-Leontief left regime, $\varepsilon(\theta) \to 1$ and $\eta_u(\theta) \to 0$, so the modified Hosios condition prescribes $\alpha_L^H \to 0$—the congestion externality vanishes and the planner would slash the shadow "wage" to approximately b, hiring aggressively. The decentralized economy, locked into $\alpha_L = 0.112$, cannot adjust. The planner shuts down vacancy posting in only one of the twenty-five productivity grid nodes (points on the solution grid, not states of a discretized Markov chain), against six for the decentralized economy.

Figure 8 isolates a representative crisis episode—the spell whose peak unemployment is closest to the median peak across all 127 crisis spells in the simulation. Before the crisis, the decentralized economy runs slightly below the planner's unemployment (blue shading in panel a): the calibrated $\alpha_L = 0.112$ lies below the efficient share $\alpha_L^H \approx 0.134$, so firms capture too much of the surplus and over-hire. When a sequence of adverse productivity shocks pushes tightness into the near-Leontief left regime, the decentralized economy shuts down vacancy posting entirely—$\theta = 0$ for three months (panel b, pink band)—while the planner, whose $\alpha_L^H(\theta) \to 0$ slashes the effective wage share, keeps tightness near 0.4 and vacancies open throughout. The result is a 9.2-percentage-point excess unemployment gap at the peak: decentralized unemployment spikes to 18.2% against the planner's 9.0%. After shutdown ends, the pattern reverses: decentralized tightness overshoots the planner as the over-hiring regime reasserts itself, and unemployment drops back below the efficient level within about ten months.

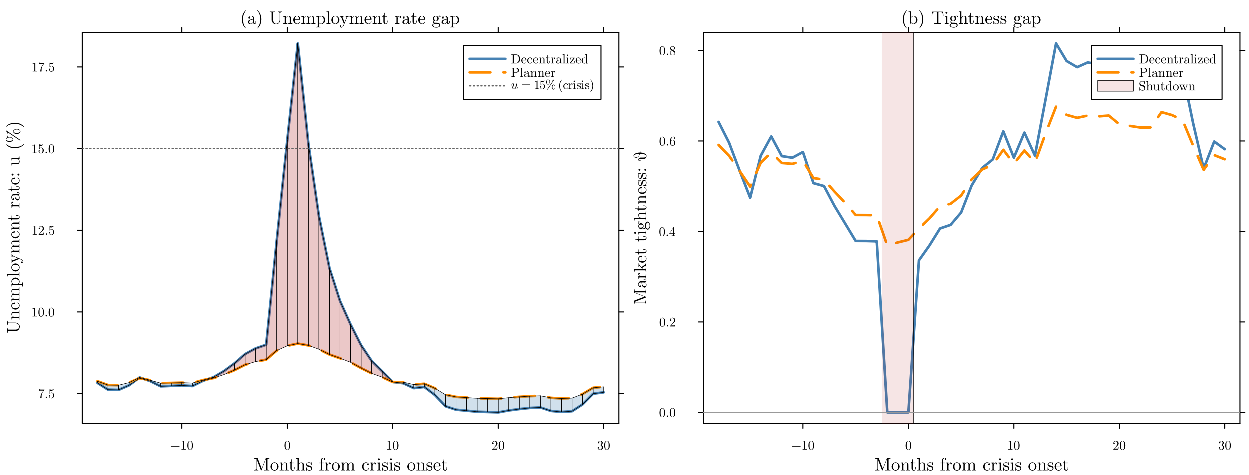

Figure 8: Representative crisis episode under the 2P-SMCD: decentralized (solid) versus planner (dashed) allocation. Panel (a): unemployment rate. Shading marks the gap (red: excess, blue: insufficient). Dotted line: 15% crisis threshold. Panel (b): market tightness. Pink band: vacancy-shutdown region ($\theta = 0$).

The HRW economy tells a very different story. Its calibrated $\alpha_L = 0.061$ sits essentially at $\alpha_L^H = 0.060$, so mean planner and decentralized unemployment nearly coincide (7.4% versus 7.6%, gap +0.19 pp). The gap splits roughly evenly between excess (55%) and insufficient (45%) unemployment, with no clear direction. In crisis episodes the gap is +6.7 pp, well below the SMCD's +11.1 pp, because HRW's smooth logistic elasticity profile never produces the sharp $\alpha_L^H \to 0$ collapse. The gap is more linearly countercyclical ($\rho(u_{\text{gap}}, z) = -0.79$ versus -0.44 for SMCD), confirming that HRW's inefficiency is spread evenly rather than concentrated in the states where it hurts most.

The upshot is that, though recalibration equalizes specific first and second moments, it cannot equalize objects that depend on curvature. The tail risk, Hosios-wedge volatility, and crisis-episode unemployment gaps all depend on p_L and θ_c, which are estimated from matching data. Appendix D.1 shows that p_L is identified only from below, and Appendix D.2 recomputes these statistics across $p_L \in [5, 20]$: the crisis gap (10.2–11.5 pp) and unemployment skewness (7.2–11.3) vary little and soften only when p_L is pushed well below its identified lower bound, so the headline results turn on the asymmetry $p_L \gg p_R$ rather than on the exact left-branch curvature.

This inefficiency relates to, but is distinct from, the model-free sufficient-statistics approach of Michaillat and Saez (2021, 2024). Michaillat and Saez (2021) derive a general formula for the *Beveridgean unemployment gap* $u - u^*$, where the efficient rate u^* equates the marginal social costs of unemployment and vacancies. Their formula requires three sufficient statistics: the elasticity of the Beveridge curve (the elasticity of v with respect to u along the uv-locus—an equilibrium object folding in separations and job creation, distinct from the matching-function elasticity $\varepsilon(\theta)$ that governs α_L^H here), the social cost of unemployment, and the cost of recruiting. The efficient tightness θ^* under this general formula need not equal one. It depends on the cost ratio and the Beveridge elasticity. Crucially, the formula that maps these sufficient statistics to u^* does not require specifying the wage-setting mechanism: the Beveridge curve, measured directly from data on u and v, already embeds whatever wage protocol generated the equilibrium.

Michaillat and Saez (2024) show that the general formula admits a powerful special case that can be immediately operationalized with readily available data. Under three assumptions—the Beveridge curve is an approximate rectangular hyperbola ($uv \approx$ const, an empirical regularity), one recruiter is required per vacancy so that a job-seeker and a vacancy are equally socially costly, and job search has no social value—the sufficient-statistics formula collapses to $u^* = \sqrt{uv}$: the market is efficient at tightness $\theta = 1$, inefficiently tight when $\theta > 1$, and inefficiently slack when $\theta < 1$.

In our environment, the fixed matching cost breaks the one-recruiter-per-vacancy symmetry, so the efficient allocation is characterized not by $\theta = 1$ but by the modified Hosios weight $\alpha_L^H(\theta)$ of Proposition 3. Our "inefficiently tight on average" is accordingly a statement about *over-hiring* ($\bar{u}^{\,\mathrm{dec}} < \bar{u}^{\,\mathrm{pln}}$ because $\alpha_L < \overline{\alpha_L^H}$), not about θ exceeding one. The novelty is the *cyclical reversal*: the bounded matching function drives $\eta_u(\theta) \to 0$ in crises, collapsing the efficient share below the fixed bargained weight and turning average over-hiring into large excess unemployment precisely when unemployment is already high.[16]

Two caveats frame these welfare statements. First, the bargained weight α_L is a calibrated convention, chosen to reproduce the wage elasticity, not an object identified from the matching data. Second, the sign-reversing gap is driven jointly by the shape of the matching function and by the *acyclicality* of the bargained share: a wage protocol delivering a procyclical worker share, such as alternating-offer bargaining, would let α_L track $\alpha_L^H(\theta)$ downward in recessions and attenuate the crisis gap. The welfare verdict is therefore

[16]The efficient tightness also differs. In the symmetric-cost framework of Michaillat and Saez (2024) it is the constant $\theta^* = 1$. In ours, recruiting costs are fixed in units of output and do not scale with z, so the planner's efficient tightness $\theta^{\mathrm{pln}}(z)$ is procyclical. Our gap thus measures the decentralized allocation against a procyclical benchmark rather than a fixed $\theta^* = 1$.

conditional on Nash bargaining with a constant weight. The more robust feature is the direction and state-dependence of the efficient share $\alpha_L^H(\theta)$, which follows from the matching technology alone.

Implementation. The planner allocation can be decentralized by a state-dependent hiring tax or subsidy levied on the firm per filled match. The relevant margin is the firm's job-creation condition (9): the firm posts vacancies until the hiring cost $\kappa + k/q(\theta)$ equals the expected value of a filled job, which depends on flow profit $z - w$. The subtlety is that Nash bargaining renegotiates the wage in response to the tax. A per-period levy T on the firm reduces the total surplus available for splitting from $z - b + k\theta$ to $z - b + k\theta - T$. Nash bargaining then lowers the negotiated wage by $\alpha_L T$, so the firm absorbs only a fraction $1 - \alpha_L$ of the tax (Appendix B.2.1). The transfer that equates the firm's effective labor cost $w + (1 - \alpha_L)T$ with the planner wage $w^{\text{pln}} = b + \alpha_L^H(\theta)(z - b + k\theta)$ is

$$T(\theta, z) = \frac{\alpha_L^H(\theta) - \alpha_L}{1 - \alpha_L} \cdot (z - b + k\theta),$$

where the factor $1/(1 - \alpha_L)$ corrects for bargaining incidence. When $T > 0$ (normal times, $\alpha_L^H \approx 0.134 > \alpha_L = 0.112$), the tax discourages over-hiring. When $T < 0$ (crises, $\alpha_L^H \to 0 < \alpha_L$), the subsidy incentivizes vacancy creation in the states where the decentralized economy shuts down. The sign flip in T mirrors the sign flip in the unemployment gap.

The relationship between matching-function shape and policy design connects to Landais et al. (2018), who characterize optimal unemployment insurance under constant matching elasticity and risk-averse workers in the Merz (1995) large-family framework. In their environment, the efficiency cost of generous benefits is lower in recessions because the congestion externality remains material throughout the cycle, making optimal UI countercyclical. The present model instead features variable matching elasticity, risk-neutral workers, and a hiring tax/subsidy rather than UI as the policy instrument. Under the SMCD's state-dependent elasticity, the congestion externality vanishes in crises ($\eta_u \to 0$), so the efficiency-side case for discouraging job creation disappears at the point where unemployment is highest—the planner instead wants a hiring subsidy.

Implementation requires observing θ in real time and knowing the matching function's shape, both of which the horse-race estimates provide. The practical nuance is that the tax/subsidy schedule $T(\theta, z)$ is nonlinear and state-dependent, requiring more information than a flat hiring credit or a fixed replacement rate.

The welfare results depend on ϕ, the fixed-cost share, set to $\phi = 0.90$ following Christiano et al. (2016). Because the modified Hosios condition compresses α_L^H toward zero as $\phi \to 1$, Appendix D.3 recalibrates the entire model at $\phi = 0.80$. The qualitative verdict is unchanged: the crisis-episode gap and unemployment skewness are virtually unchanged, and the low calibrated bargaining weight is confirmed to come from the high fixed-cost share rather than from any intrinsic feature of the matching function.

We have quantified the accuracy of the global solution, and of the crisis statistics in particular. The Euler residuals of the job-creation condition average $10^{-5.8}$ across the simulated ergodic distribution. They are largest near the shutdown corner, where crises occur, but even the worst case stays below 10^{-3}—an error under 0.1%. The skewness and crisis gap are stable under refinement of the productivity grid to 200 nodes, and the solution passes a den Haan–Marcet test of whether its one-step-ahead forecast errors can be predicted from current information, as they could not be if the policy were exact. Appendix D.4 reports the full accuracy suite, including the convergence of the tail statistics with the length of the simulation.

7. Discussion: The Continuous-Time Objection

One might object that the probability-bound violations motivating the SMCD are an artifact of discrete time and disappear in continuous-time formulations. In a standard continuous-time DMP model, the matching function $m(u, v)$ specifies a *flow rate* of matches per unit time, and the associated job-finding and vacancy-filling objects $f(\theta) = m(1, \theta)$ and $q(\theta) = m(1, \theta)/\theta$ are hazard rates rather than probabilities. Hazard rates are non-negative but not bounded above by one: a Poisson matching process can deliver $f(\theta) > 1$ per unit time without logical contradiction, since the implied per-period probability $1 - e^{-f(\theta)\Delta t}$ is in $(0, 1)$ for any finite $\Delta t > 0$. On this view, the CD function $f(\theta) = A\theta^{1-\alpha}$ is a valid hazard-rate specification for all θ, and the violations documented in Table 1 are an artifact of applying a continuous-time object directly as a monthly probability.

The objection is technically correct but overstates what continuous time resolves, and its surviving force turns out to bear entirely on the vacancy-filling rate q at low tightness: even under the exact hazard-to-probability mapping the job-finding rate f is bounded, but q is not. The first difficulty is empirical. Matching functions are estimated by confronting model-implied objects with data at monthly frequency—the standard practice from Shimer (2005) to Hagedorn and Manovskii (2008). A model that predicts $f(\theta) > 1$ at realized tightness values is misspecified at the calibration frequency, regardless of whether it is written in continuous or discrete time.

The second difficulty cuts against the objection on its own terms. The correct link between a continuous-time hazard $\lambda_f(\theta)$ and the monthly probability is $f_t = 1 - e^{-\lambda_f(\theta)}$, not the first-order approximation $f_t \approx \lambda_f(\theta)$. Applied to the CD hazard $\lambda_f(\theta) = A\theta^{1-\alpha}$, this yields

$$f_t(\theta) = 1 - \exp\Big(-A\theta^{1-\alpha}\Big), \tag{13}$$

which lies in $(0,1)$ for all θ and has a variable elasticity $\varepsilon(\theta) = A(1-\alpha)\theta^{1-\alpha}/\Big(\exp(A\theta^{1-\alpha})-1\Big)$ that declines with θ. The practitioner who applies this mapping consistently has abandoned the log-linear matching function in favor of a bounded, variable-elasticity specification sharing the core features of the SMCD. But (13) bounds only f, not q: since $q(\theta) = f(\theta)/\theta \to \infty$ as $\theta \to 0$, the exponential specification still predicts $q > 1$ below a threshold that includes the deepest recession observations. At the COVID trough $\theta \approx 0.15$, (13) with $\hat{A} = 0.45$ and $\hat{\alpha} = 0.71$ implies $q \approx 1.53$. The SMCD bounds both $f \leq s$ and $q \leq sa$ simultaneously.

Ultimately, the strongest case for bounded matching functions is quantitative. Section 6 shows that the near-Leontief branch of the 2P-SMCD—a direct consequence of the vacancy-filling bound—drives unemployment skewness (10.7 versus 1.8 under HRW), fat-tailed crisis risk, and a Hosios efficiency wedge that flips sign across the cycle. The quantitative implications for unemployment dynamics and welfare are large enough to justify the additional parameter.

8. Conclusion

The SMCD matching function resolves a tension between the two standard specifications in search theory: Cobb-Douglas fits the data well but predicts impossible vacancy-filling probabilities in recessions, while HRW respects the bounds but cannot calibrate jointly to a target job-finding rate and elasticity. The SMCD does both, with only three parameters and closed-form calibration formulas.

A QLR sup-Wald test identifies a structural break in the U.S. matching relationship at October 2008. Across both subsamples the two-plateau SMCD dominates HRW and closes most of Cobb-Douglas's residual RMSE advantage, which survives only in the recession months where Cobb-Douglas predicts an impossible vacancy-filling probability. Because the vacancy-filling bound binds at the estimated optimum, the 2P-SMCD collapses to the three-parameter closed form of (5), which practitioners can adopt directly using the estimates in Table 6 for their sample period, with no re-estimation required. Embedding the estimated shape in a Diamond–Mortensen–Pissarides economy with the two-part hiring cost of Pissarides (2009) makes the Hosios condition state-dependent through both the matching elasticity and the fixed-cost share. Calibrated to the same three targets, the SMCD economy is inefficiently tight on average, echoing Michaillat and Saez (2024), but the gap reverses to +11.1 pp of excess unemployment in crises, against +6.7 pp under HRW. The same near-Leontief branch preserves the vacancy-shutdown channel of Petrosky-Nadeau and Zhang (2017) under joint calibration.

Two extensions are natural: an out-of-sample test on the post-2008 tightness range, which spans both limiting regimes, and a model with risk-averse agents that revisits the optimal-UI question of Landais et al. (2018) against the vanishing congestion externality documented here.

References

Anderson, P.M., Burgess, S.M., 2000. Empirical matching functions: Estimation and interpretation using state-level data. Review of Economics and Statistics 82, 93–102. doi:10.1162/003465300558597.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856. URL: https://doi.org, doi:10.2307/2951764.

Angelis, T., Bramoullé, Y., 2026. Matching functions and networks. Working Paper.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78. doi:10.2307/2998540.

Bai, Y., Ríos-Rull, J.V., Storesletten, K., 2025. Demand shocks as technology shocks. Review of Economic Studies 00, 1–35. doi:10.1093/restud/rdaf045.

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

Barnichon, R., Figura, A., 2015. Labor market heterogeneity and the aggregate matching function. American Economic Journal: Macroeconomics 7, 222–249.

Borowczyk-Martins, D., Jolivet, G., Postel-Vinay, F., 2013. Accounting for endogeneity in matching function estimation. Review of Economic Dynamics 16, 440–451. doi:10.1016/j.red.2012.12.001.

Christiano, L.J., Eichenbaum, M.S., Trabandt, M., 2016. Unemployment and business cycles. Econometrica 84, 1523–1569. doi:10.3982/ECTA11776.

Den Haan, W.J., Marcet, A., 1994. Accuracy in simulations. The Review of Economic Studies 61, 3–17. doi:10.2307/2297873.

Diamond, P.A., 1982. Aggregate demand management in search equilibrium. Journal of political Economy 90, 881–894.

Elsby, M.W.L., Michaels, R., Ratner, D., 2015. The Beveridge curve: A survey. Journal of Economic Literature 53, 571–630. doi:10.1257/jel.53.3.571.

den Haan, W.J., Ramey, G., Watson, J., 2000. Job destruction and propagation of shocks. American Economic Review 90, 482–498.

Hagedorn, M., Manovskii, I., 2008. The cyclical behavior of equilibrium unemployment and vacancies revisited. American Economic Review 98, 1692–1706. URL: http://www.aeaweb.org/articles?id=10.1257/aer.98.4.1692, doi:10.1257/aer.98.4.1692.

Hall, R.E., Milgrom, P.R., 2008. The limited influence of unemployment on the wage bargain. American economic review 98, 1653–74.

Hosios, A.J., 1990. On the efficiency of matching and related models of search and unemployment. The Review of Economic Studies 57, 279–298. doi:10.2307/2297382.

Kiyotaki, N., Wright, R., 1993. A search-theoretic approach to monetary economics. American Economic Review 83, 63–77.

Lagos, R., Wright, R., 2005. A unified framework for monetary theory and policy analysis. Journal of Political Economy 113, 463–484. doi:10.1086/429804.

Landais, C., Michaillat, P., Saez, E., 2018. A macroeconomic approach to optimal unemployment insurance: Applications. American Economic Review 108, 1982–2029. doi:10.1257/aer.20160055.

Lange, F., Papageorgiou, T., 2020. Beyond Cobb-Douglas: Flexibly Estimating Matching Functions with Unobserved Matching Efficiency. Working Paper 26972. National Bureau of Economic Research. doi:10.3386/w26972.

Lubik, T.A., 2013. The Shifting and Twisting Beveridge Curve: An Aggregate Perspective. Working Paper 13-16. Federal Reserve Bank of Richmond. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2335720.

Mangin, S., Julien, B., 2021. Efficiency in search and matching models: A generalized Hosios condition. Journal of Economic Theory 193, 105208. doi:10.1016/j.jet.2021.105208.

Merz, M., 1995. Search in the labor market and the real business cycle. Journal of monetary Economics 36, 269–300.

Michaillat, P., Saez, E., 2021. Beveridgean unemployment gap. Journal of Public Economics Plus 2, 100009. doi:10.1016/j.pubecp.2021.100009.

Michaillat, P., Saez, E., 2024. $u^* = \sqrt{uv}$: The Full-Employment Rate of Unemployment in the United States. Working Paper 2206.13012. arXiv. doi:10.48550/arXiv.2206.13012.

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

Petrongolo, B., Pissarides, C.A., 2001. Looking into the black box: A survey of the matching function. Journal of Economic literature 39, 390–431.

Petrosky-Nadeau, N., Wasmer, E., 2017. Labor, Credit, and Goods Markets: The Macroeconomics of Search and Unemployment. MIT Press.

Petrosky-Nadeau, N., Zhang, L., 2017. Solving the Diamond–Mortensen–Pissarides model accurately. Quantitative Economics 8, 611–650. doi:10.3982/QE452.

Petrosky-Nadeau, N., Zhang, L., 2021. Unemployment crises. Journal of Monetary Economics 117, 335–353. doi:10.1016/j.jmoneco.2020.01.009.

Pissarides, C.A., 1985. Short-run equilibrium dynamics of unemployment, vacancies, and real wages. American Economic Review 75, 676–690.

Pissarides, C.A., 2000. Equilibrium Unemployment Theory. 2nd ed., The MIT Press, Cambridge, Massachusetts.

Pissarides, C.A., 2009. The unemployment volatility puzzle: Is wage stickiness the answer? Econometrica 77, 1339–1369. doi:10.3982/ECTA7562.

Şahin, A., Song, J., Topa, G., Violante, G.L., 2014. Mismatch unemployment. American Economic Review 104, 3529–3564.

Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review 95, 25–49.

Shimer, R., 2012. Reassessing the ins and outs of unemployment. Review of Economic Dynamics 15, 127–148.

Silva, M.R., Urias, M., 2025. Productive demand and sectoral capacity utilization. Economic Modelling 153, 107317. doi:10.1016/j.econmod.2025.107317.

Stevens, M., 2007. New microfoundations for the aggregate matching function. International Economic Review 48, 847–868. doi:10.1111/j.1468-2354.2007.00447.x.

Online Appendix

For Online Publication

Appendix A. Additional Figures

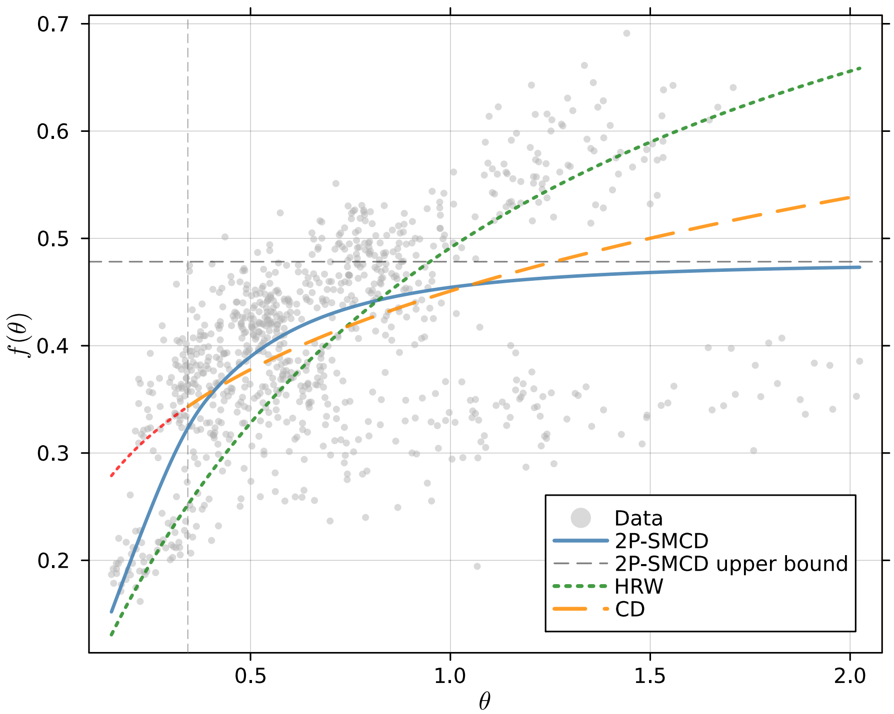

Figure A.1: Full-sample horse race: job-finding rate $f(\theta)$ for 2P-SMCD (solid), Cobb-Douglas (dashed), and HRW (dotted), 1951–2025. Grey scatter: data. Dashed horizontal line: 2P-SMCD upper bound s_R. Red dotted: CD where fitted $q > 1$. Vertical line: $\theta_{\min}$.

Table 5 splits the sample at the structural break. Estimated instead on the pooled sample, the 2P-SMCD attains a lower RMSE than Cobb-Douglas both overall (0.0772 versus 0.0794) and on the valid domain $\theta > \theta_{\min}$ where CD respects $q \leq 1$ (0.0784 versus 0.0796), while Cobb-Douglas implies $q > 1$ in 13.5% of months. Pooling the two regimes reverses the pre-break ranking of Table 5. The two-plateau form absorbs the level shift across the break that a rigid power law cannot, so the 2P-SMCD fits better even on the domain where Cobb-Douglas is a valid probability model.

Both the 2P-SMCD and Cobb-Douglas undershoot the high-job-finding observations at high tightness in the northeast of the scatter. The 2P-SMCD misses them by the widest margin, because its bounded right branch saturates at the ceiling s_R while a power law keeps rising. In return the 2P-SMCD tracks the dense central cluster more closely, which is the source of its lower overall RMSE. The broader lesson is that no smooth, monotonic $f(\theta)$ fits the pooled data well. The structural break superimposes two matching regimes with different efficiency, so at a given tightness the pre-break and post-break job-finding rates differ, and a single monotonic curve cannot pass through both.

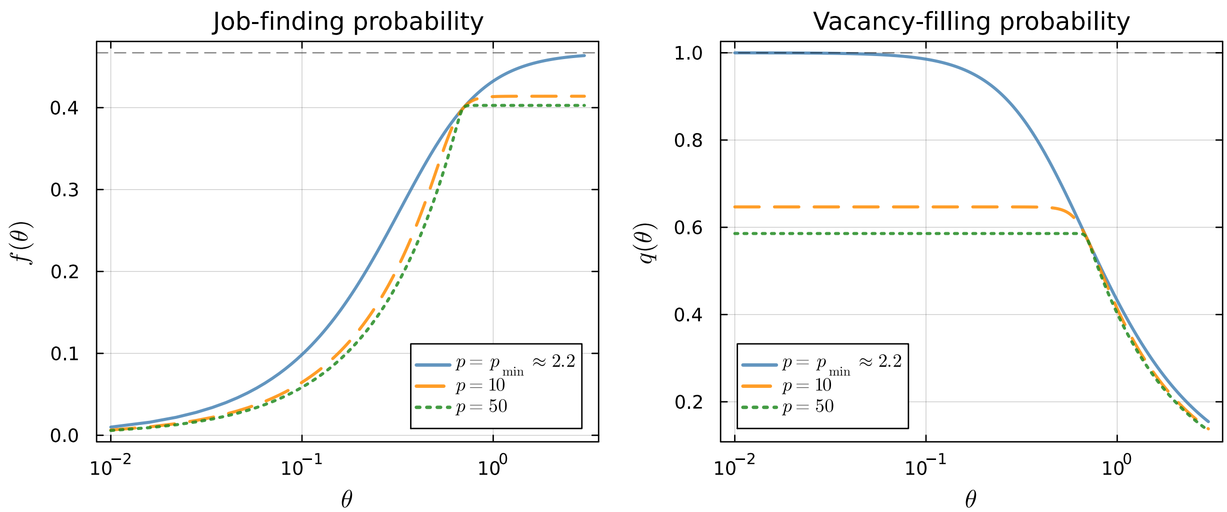

Figure A.2: SMCD job-finding rate (left) and vacancy-filling rate (right) for three values of p. Calibration: $(\theta^*, f^*, \varepsilon^*) = (0.70, 0.40, 0.29)$; (s, a) recalibrated for each p. Dashed lines: asymptotes $f = s$ and $q = sa$ at $p = p_{\min}$.

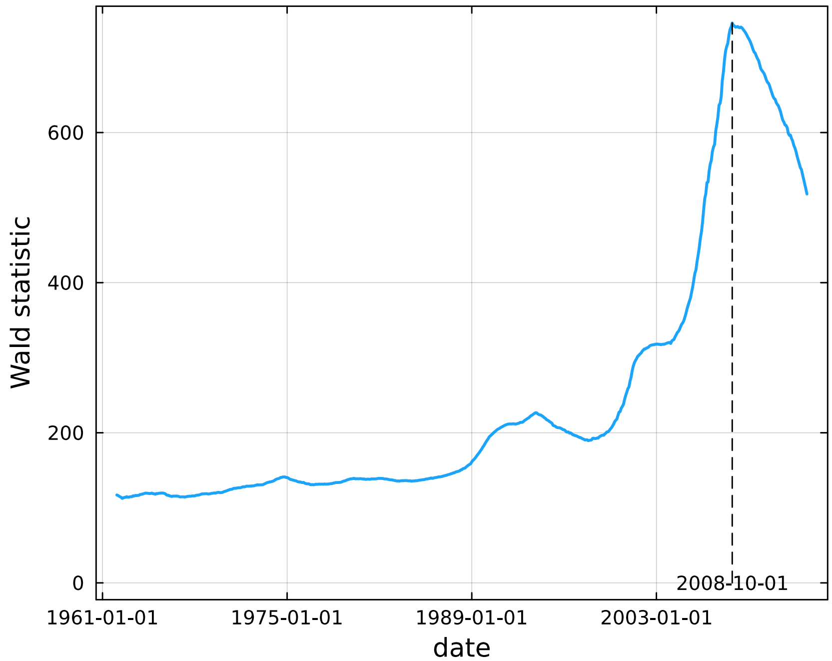

Figure A.3: QLR sup-Wald statistic $W(\tau)$ for a structural break in the log-log matching relationship $\log f_t = \alpha + \beta \log \theta_t + u_t$. Vertical dashed line: estimated break date, October 2008. $\sup W = 745.97$. Bootstrap p-value = 0.000.

Appendix B. Proofs and Derivations

Appendix B.1. Concavity Proof

Let $z = a\theta$ and define $h(z) = z(1 + z^p)^{-1/p}$. As $f(\theta) = sh(a\theta)$, $f''(\theta) = sa^2 h''(a\theta)$, so it suffices to show $h''(z) \leq 0$ for $z > 0$. Calculations yield

$$h'(z) = (1 + z^p)^{-1/p-1}$$
$$h''(z) = -(p+1)z^{p-1}(1 + z^p)^{-2-1/p} \leq 0 \quad \text{for} \quad z > 0, p > 0.$$

Appendix B.2. The Modified Hosios Condition with Fixed Matching Costs

This appendix derives Proposition 3 from the dynamic planner's problem of the stochastic environment of Section 6.1, following the planner-versus-decentralization logic of Hosios (1990) and Pissarides (2000, ch. 8) extended to the two-part hiring cost. The efficiency condition is a *within-period* arbitrage that holds separately at each productivity state z. The deterministic steady state ($z_t = \bar{z}$) is the special case, and the continuation values a dynamic derivation carries cancel from the condition. This is why solving the planner's allocation state by state with $\alpha_L^H(\theta)$—as Section 6.4 does—is exact, not a steady-state approximation.

Decentralization. Let $S(z)$ denote the total surplus of a filled match, $S(z) = J(z) + \Big(W(z) - U\Big)$, the sum of the firm's job value (7) and the worker's net value. Generalized Nash bargaining with worker weight α_L splits it, $W(z) - U = \alpha_L S(z)$ and

$$J(z) = (1 - \alpha_L)\, S(z). \tag{B.1}$$

The fixed matching cost κ is borne by the firm at the moment of formation and is not part of the split—consistent with the Nash wage (10), in which κ does not appear. Whenever vacancies are posted, free entry drives the value of a vacancy to zero: a vacancy costs k per period and fills at the private rate $q(\theta)$, delivering the match net of the fixed cost, so

$$\frac{k}{q(\theta(z))} = J(z) - \kappa \quad \iff \quad J(z) = \kappa + \frac{k}{q(\theta(z))}. \tag{B.2}$$

Both (B.1) and (B.2) hold state by state.

Planner. A utilitarian planner chooses tightness to maximize the present value of output net of the flow value of nonemployment and of hiring costs,

$$\max_{\{\theta_t \geq 0\}} \; \mathbb{E}_0 \sum_{t \geq 0} \beta^t \Big[z_t(1 - u_t) + b\, u_t - k\, \theta_t u_t - \kappa\, u_t f(\theta_t) \Big], \qquad u_{t+1} = u_t + \delta(1 - u_t) - u_t f(\theta_t).$$

Let $\lambda(z)$ be the social value of an additional filled job—the planner's counterpart of $J(z)$. An extra vacancy raises the flow of matches not at the private rate $q(\theta)$ but at the social rate $\partial m / \partial v = \varepsilon(\theta)\, q(\theta)$, where $\varepsilon(\theta) = \partial \ln m / \partial \ln v$, so the planner's vacancy first-order condition equates the flow cost to the social matching rate times the net social value, $k = \varepsilon(\theta)\, q(\theta)\, (\lambda(z) - \kappa)$, i.e.

$$\lambda(z) = \kappa + \frac{k}{\varepsilon(\theta(z))\, q(\theta(z))}, \tag{B.3}$$

which again holds state by state. Like the private surplus, the social value obeys a forward recursion, $\lambda(z) = (z - b) + \beta(1 - \delta)\, \mathbb{E}[\lambda(z') \mid z]$, with the *same* discount factor $\beta(1 - \delta)$ and the same flow surplus $z - b$ that drive $S(z)$. The congestion externality enters only through the matching rate $\varepsilon(\theta) q(\theta)$ in (B.3), not through the continuation.

Efficiency. The decentralized and planner allocations coincide at every state when $S(z) = \lambda(z)$: the two surpluses then satisfy an identical recursion, so equality at every z is consistent and both economies choose the same $\theta(z)$. Imposing $S = \lambda$ and combining the within-period relations (B.1), (B.2), and (B.3),

$$(1-\alpha_L)\left(\kappa + \frac{k}{\varepsilon\, q}\right) = \kappa + \frac{k}{q}.$$

Writing $x \equiv k/q(\theta)$ and solving for α_L,

$$\alpha_L = \frac{x(1-\varepsilon)}{\varepsilon\kappa + x}.$$

Dividing numerator and denominator by $\kappa + x$ and using the fixed-cost share $\phi(\theta) = \kappa/(\kappa + x)$ from (6), together with $\eta_u(\theta) = 1 - \varepsilon(\theta)$, gives

$$\alpha_L^H(\theta) = \frac{(1-\phi(\theta))\,\eta_u(\theta)}{1-\phi(\theta)\,\eta_u(\theta)},$$

which is (12). The continuation values have dropped out: efficiency is the within-period requirement that the surplus split neutralize the congestion externality, imposed at each θ separately, so the same $\alpha_L^H(\theta)$ characterizes both the deterministic steady state and every state of the dynamic economy. At $\phi = 0$ (purely flow costs, $\kappa = 0$) it returns the standard condition $\alpha_L = \eta_u(\theta) = 1 - \varepsilon(\theta)$. At $\phi \to 1$ it tends to zero. Only the flow cost $x = k/q(\theta)$ responds to congestion, so the fixed matching cost dilutes the externality that Nash bargaining must internalize by the factor $1 - \phi(\theta)$.

The vacancy-shutdown corner. Because (B.3) and (B.2) are within-period arbitrage conditions, the non-negativity constraint $\theta \geq 0$ leaves the interior efficiency condition unchanged: it only fixes the states in which posting ceases. The corner binds for the planner and for the market at (generally different) thresholds in $\lambda(z)$ and $J(z)$. Wherever both post, (12) characterizes efficiency, and the state-dependent allocation of Section 6.4 imposes $\alpha_L^H(\theta)$ throughout. □

Appendix B.2.1. Decentralizing the planner allocation

This subsection derives the state-dependent hiring tax or subsidy $T(\theta, z)$ that makes the decentralized equilibrium replicate the planner allocation of Proposition 3. The key subtlety is that Nash bargaining renegotiates the wage in response to any tax levied on the firm, so only a fraction of the tax is borne by the firm. The required instrument must compensate for this pass-through.

The Nash wage (10) splits the per-period match surplus $\Sigma \equiv z - b + k\theta$ between worker and firm:

$$w = b + \alpha_L \, \Sigma, \qquad z - w = (1 - \alpha_L) \, \Sigma.$$

The fixed matching cost κ is sunk at the point of bargaining and does not appear in Σ: it enters only the free-entry condition (8). Now suppose the government levies a state-dependent per-period transfer $T(\theta, z)$ on the firm ($T > 0$: tax; $T < 0$: subsidy). Nash bargaining splits the reduced surplus $\Sigma - T$, giving

$$w^\tau = b + \alpha_L \, (\Sigma - T) = w - \alpha_L \, T.$$

The firm's total per-period labor cost—wage plus tax—is therefore

$$w^\tau + T = w + (1 - \alpha_L) \, T : \tag{B.4}$$

Nash renegotiation absorbs a fraction α_L of the tax, so only $(1 - \alpha_L)T$ falls on the firm. The planner allocation is replicated when the firm faces the effective wage it would pay under $\alpha_L^H(\theta)$, i.e. $w^\tau + T = b + \alpha_L^H(\theta) \, \Sigma$. Substituting (B.4) and solving,

$$T(\theta, z) = \frac{\alpha_L^H(\theta) - \alpha_L}{1 - \alpha_L} \; \cdot \; (z - b + k\theta).$$

When $\alpha_L^H(\theta) > \alpha_L$ (the decentralized economy over-hires), $T > 0$: the tax discourages vacancy posting. When $\alpha_L^H(\theta) < \alpha_L$ (the decentralized economy under-hires), $T < 0$: the subsidy incentivizes vacancy creation in the states where the congestion externality vanishes. $\square$

Appendix C. Global Solution and Estimation

This appendix details the two nested numerical steps summarized in Section 6.3: an inner *global* solution of the model by time iteration on the job-creation condition, and an outer estimation that matches moments of the model's simulated stationary distribution to their empirical counterparts. The two matching functions are solved and estimated identically. Table C.1 collects the numerical settings.

Appendix C.1. Time iteration

Equilibrium tightness solves the job-creation condition (9), a fixed-point problem in the value of a filled job J as a function of the productivity state z. Rather than discretize $\log z$ on a Markov chain, we work with the expected continuation value $E(z) \equiv \mathbb{E}[J(z') \mid z]$ directly, evaluating the conditional expectation by Gauss–Hermite quadrature over the continuous log-normal innovation distribution. This eliminates the discretization staircase in policy functions that a finite grid introduces, which would otherwise distort the tail behavior of the simulated unemployment distribution. The iteration variable $E(z)$ is represented on a 25-point grid in z. Conditional expectations at off-grid productivity realizations during simulation are obtained by linear interpolation. Time iteration works directly with the equilibrium free-entry condition rather than a value function. From an initial guess, each pass performs the following.

1. **Initialize** $E(z)$ at the deterministic steady-state continuation value.
2. **Free-entry inversion.** At each grid point, invert $q(\theta_i) = k/(E_i + \kappa_{\text{net}} - \kappa)$ for tightness, where κ_{net} accounts for the next-period fixed cost. The inverse q^{-1} is analytic for both the two-plateau SMCD and the HRW. Shutdown states are handled by the ramp introduced below.
3. **Flow profit.** Substitute the Nash wage (10) to obtain $z_i - w_i = (1-\alpha_L)(z_i - b) - \alpha_L k\theta_i$.
4. **Howard improvement.** Take $K = 15$ partial policy-evaluation steps at fixed θ, iterating $E(z) \leftarrow \sum_m w_m \left[(z'_m - w'_m) + (1-\delta)(\kappa + k/q(\theta'_m))\right]$, where the sum runs over Gauss–Hermite nodes z'_m with weights w_m and θ'_m is the interpolated policy at z'_m.
5. **Damped update and convergence.** Compute $\text{err} = \max_i |\Delta E_i|$ and update $E \leftarrow \omega\, E^{\text{new}} + (1-\omega)\, E$ with the adaptive ω introduced below. Return to step 2 until $\text{err} < 10^{-7}$ or 2000 passes.

Working from the free-entry condition inherits the fast convergence of policy iteration and never iterates a value function.

Table C.1: Numerical implementation.

Setting	Value
Grid points N; GH quadrature nodes	25; 7
Productivity AR(1) (ρ_z, σ_z), monthly	(0.911, 0.0053)
Discount β; separation δ	$e^{-0.04/12}$; 0.035
Fixed-cost share ϕ	0.90
Tightness bounds $[\theta_{\text{lo}}, \theta_{\text{hi}}]$	$[10^{-3},\ 30]$
Howard steps K	15
Outer damping ω	init 0.3, range [0.02, 0.9]
Convergence tolerance; max passes	10^{-7}; 2000
Homotopy on $\sigma(u)/\sigma(z)$	$5.0 \to 10.8$, 6 steps
Inner solver (b, k)	trust region, ftol 10^{-4}, ≤ 300 iter
Outer search α_L	false position, $[0.05, 0.30]$, tol 0.008, ≤ 6
Simulation length (moments; tail objects)	200,000; 5,000,000 months
Burn-in; seed	1,200; 123

Appendix C.2. The bounded-q corner

The cap $q \leq \bar{q}$ makes the inverse q^{-1} steep near the shutdown threshold $\bar{J} \equiv \kappa + k/\bar{q}$: in the near-Leontief left regime q is nearly flat, so a vanishing change in J moves θ sharply away from zero, and an undamped iteration oscillates in a limit cycle at the boundary. Two devices tame this while altering the iteration path but not the fixed point. First, *during* the iteration the hard shutdown is replaced by a narrow smooth ramp: θ is scaled linearly from 0 to its full value over the band $[\bar{J}, \bar{J} + \delta_s]$, with $\delta_s = \text{clamp}\Big(0.15\,(J_{\text{ss}} - \bar{J}),\, 10^{-4},\, 2k/\bar{q}\Big)$ a small fraction of the (narrow) gap between the steady-state J and $\bar{J}$, so the steady state is never attenuated. The reported policy uses the exact shutdown $\theta = 0$ for $J \leq \bar{J}$. Second, the outer damping weight is adaptive: ω starts at 0.3 and is scaled by 1.1 toward 0.9 when err falls, or by 0.5 toward a floor of 0.02 when it rises. Convergence is judged in absolute terms, $\max_i |\Delta J_i|$, because the thin surplus at the calibration drives J toward zero in slack states, where a relative criterion is unreliable.

Appendix C.3. Estimation and homotopy

For a candidate (b, α_L, k) we fix κ from the fixed-cost share $\phi = 0.90$, solve the model as above, simulate a long panel from the continuous AR(1) process using the solved policy (with interpolated $\theta(z)$ at off-grid z values), and compute the target moments on the resulting stationary distribution as HP-filtered proportional deviations from the mean. The three data moments are matched in a nested loop that exploits their mapping to parameters

(Section 6.3):

- an *outer* bracketed (false-position) search chooses α_L over $[0.05, 0.30]$ so that the simulated wage elasticity equals its target, to a tolerance of 0.008 in at most six evaluations; and

- for each α_L, an *inner* trust-region solve chooses (b, k)—reparametrized as $b \in [0.25, 0.995]$ (logistic) and $k = 0.005 + e^x$—so that simulated $\sigma(u)/\sigma(z)$ and $\bar{\theta}$ hit their targets, with function tolerance 10^{-4} and at most 300 iterations.

The inner solve is reached by *homotopy*: the unemployment-volatility target $\sigma(u)/\sigma(z)$ is raised from 5.0 to its data value 10.8 in six equal steps, and the trust-region solve at each step is warm-started from the previous step's (b, k). Continuation is needed because the high-volatility target is a thin-surplus regime in which the job-creation operator is stiff—its contraction modulus approaches one and many states sit near the shutdown corner—so a cold solve at the final target rarely converges. Calibration and the reported second moments use a 200,000-month simulation (burn-in 1,200, seed 123), so the calibrated and reported $\sigma(u)/\sigma(z)$ are computed on the same draw. The tail statistics—level skewness, crisis frequency, and the crisis gap—are high-variance and are instead reported from a 5,000,000-month simulation at which they converge (Appendix D.4). The productivity innovations are held fixed across evaluations, so the objective is smooth in the parameters. Increasing the Gauss–Hermite node count from 7 to 15 changes the reported moments by less than 1%. We use 7 throughout.

Appendix C.4. Why a global solution

A perturbation or log-linear approximation around the deterministic steady state is blind to precisely the features this exercise studies. The shutdown corner is non-analytic, and the convex tightness policy in the near-Leontief left regime—the source of the skewed unemployment distribution and the countercyclical Hosios wedge—is a global object that a local expansion smooths away. Matching those moments therefore *requires* the global solution. A perturbation-based method would target objects the model cannot, by construction, produce.

The exercise follows the methodological principle of Petrosky-Nadeau and Zhang (2021), who show that "steady state relations hold poorly in simulations" of globally solved DMP models and therefore calibrate their recruiting-cost parameters to moments of the simulated ergodic distribution rather than to steady-state targets. We adopt the same discipline: the moments we match, and the results we report, are computed on the globally simulated stationary distribution, so the same nonlinear dynamics that generate the skewness, the crisis frequency, and the Hosios-wedge volatility also discipline the parameters. What distinguishes the present application is the separation of identification: the matching-function shape is pinned by the horse-race estimates of Section 5.4, and only the three structural parameters (b, α_L, k)—which the matching data cannot speak to—are estimated inside the model.

Appendix D. Identification, Robustness, and Numerical Accuracy

This appendix establishes that the quantitative results of Section 6 are artifacts of neither the estimated parameters nor the numerical solution. It documents parameter identification (Appendix D.1), robustness of the crisis objects to the two parameters the data do not pin down—the left-branch curvature p_L (Appendix D.2) and the fixed-cost share ϕ (Appendix D.3)—and the accuracy of the global solution (Appendix D.4).

Appendix D.1. Parameter identification

This subsection documents the estimator and the identification of the four 2P-SMCD parameters $(\theta_c, f_c, p_L, p_R)$.

Estimator. The 2P-SMCD is fit by nonlinear least squares in levels. We parametrize the junction level as a fraction of its probability ceiling, $f_c = c \cdot \Lambda(\varsigma)$ with $c = \min\left(2^{-1/p_R},\ \theta_c\, 2^{-1/p_L}\right)$ and Λ the logistic function. The advantage is that the two probability-bound inequalities ($f \leq 1$ and $q \leq 1$) collapse into a single unconstrained, smooth coordinate. The ceiling c is recomputed from (θ_c, p_L, p_R) at each evaluation, and Λ maps any $\varsigma \in \mathbb{R}$ onto $(0, c)$. The bounds therefore hold by construction, an ordinary Gauss–Newton routine applies with no inequality constraints or clamping, and a binding bound (here the vacancy-filling ceiling $q_{\max} = 1$) is approached smoothly as $\varsigma \to \infty$ rather than stalling on a boundary. The objective is then the pure sum of squared residuals, with no penalty term.

Identification of the curvature p_L. The left-branch curvature p_L is identified only from below. Once the branch is steep enough for near-Leontief behavior over the observed range of tightness, further steepening changes the fitted function negligibly, so the objective is monotone and asymptotically flat in p_L. The middle panels of Figures D.1 and D.2 display this shape: fixing p_L and re-optimizing the remaining parameters, the profile falls steeply as p_L rises from its lower bound and then flattens. The 95% profile bound is crossed only from below, at $p_L \geq 10.4$ pre-break and $p_L \geq 5.8$ post-break, with no upper crossing. What the data identify is that the left branch is an order of magnitude steeper than the right ($p_L \gg p_R$), the asymmetry that drives the near-Leontief crisis mechanism of Section 6.

A free NLS therefore drives p_L essentially to the hard-min Leontief limit, which contradicts the smoothness that motivates the soft-min form. We resolve this by calibrating p_L to the smoothest value consistent with the data. On the grid $p_L \in \{2, 2.5, \ldots, 40\}$ we record the best attainable fit $\text{RMSE}_{\min}$ and set p_L^* to the smallest p_L whose RMSE lies within $\tau = 0.5\%$ of it. This selects $p_L^* = 12.5$ pre-break and 10.5 post-break. The choice is robust to τ: at $\tau = 0.5\%, 1\%, 2\%$ the rule returns $12.5, 9.5, 6.5$ pre-break and $10.5, 7.5, 5.0$ post-break. Section Appendix D.2 confirms that the quantitative results of Section 6 are stable across this range. Because the objective is flat near the top of the grid, the exact value returned is mildly sensitive to the solver's starting point, but every admissible value places the left branch an order of magnitude steeper than the right, which is all the crisis mechanism requires.

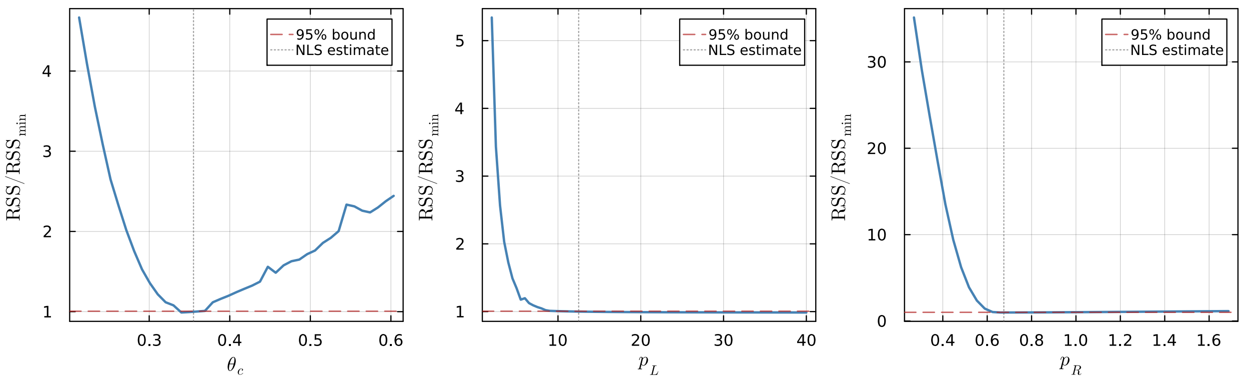

Figure D.1: Profile NLS objective for the 2P-SMCD, pre-break subsample. Each panel fixes one parameter and re-optimizes over the remaining three. p_L is profiled on a fixed grid, the others on a neighborhood of the estimate. The junction level f_c is omitted (see text). Dashed red: 95% confidence bound $1+\chi^2_1(0.95)/n$. Dotted grey: NLS estimate, and the calibrated value in the p_L panel.

Identification of the remaining parameters. Conditional on p_L, the crossover tightness θ_c and the right-branch curvature p_R are sharply identified. Their profiles (left and right panels of Figures D.1 and D.2) are steep, roughly quadratic valleys with tight 95% confidence intervals: $\theta_c \in [0.34, 0.36]$ and $p_R \in [0.65, 0.90]$ pre-break, $\theta_c \in [0.17, 0.22]$ and $p_R \in [0.64, 1.01]$ post-break. The junction level f_c is not a separate free direction, so it is not profiled. The vacancy-filling bound binds at the optimum in both subsamples, so $q_{\max} = 1$ and f_c sits exactly on its ceiling $\theta_c\, 2^{-1/p_L}$. It is determined jointly by θ_c and p_L along the binding cap. This is the same fact the horse race reports as $q_{\max} = 1$: the data want the vacancy-filling rate to reach its bound in slack markets, which is precisely why unconstrained Cobb-Douglas overshoots $q > 1$ there.

Appendix D.2. Robustness of the crisis objects to p_L

Because p_L is identified only from below, we verify that the headline quantitative results of Section 6 do not hinge on its exact value. Table D.1 recomputes the crisis objects across $p_L \in [5, 20]$, spanning the calibrated estimate and the region identified from below (lower bound $p_L \approx 10.4$), holding the remaining matching-function parameters (θ_c, f_c, p_R) at their pre-break NLS estimates and the structural parameters (b, α_L, k, ϕ) at their calibrated values (Table 7). Only p_L varies. Nothing is re-estimated.

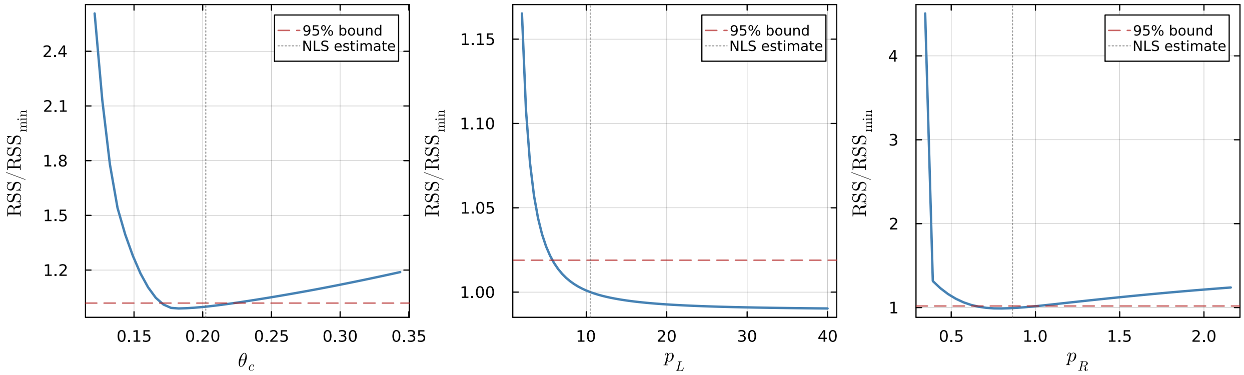

Figure D.2: Profile NLS objective for the 2P-SMCD, post-break subsample. Layout as in Figure D.1.

Table D.1: Crisis objects across a range of p_L around the calibrated estimate (2P-SMCD, pre-break).

p_L	Mean crisis gap (pp)	Skewness of u	Crisis frequency (%)	$\sigma(u)/\sigma(z)$
5.0	10.2	7.2	0.09	8.4
8.0	10.7	9.4	0.16	9.8
10.5	10.9	10.3	0.19	10.4
12.5 (estimate)	11.1	10.7	0.22	10.8
16.0	11.3	11.2	0.25	11.3
20.0	11.5	11.3	0.28	11.6

Note: Only p_L varies. (θ_c, f_c, p_R) and (b, α_L, k, ϕ) are held fixed. Tail objects (crisis gap, skewness, crisis frequency) from a 5,000,000-month simulation as in Section 6; $\sigma(u)/\sigma(z)$ from the 200,000-month calibration draw. $\sigma(u)/\sigma(z)$ is an untargeted outcome here, since the structural parameters are not recalibrated as p_L moves. It is targeted only at the estimate.

Above the identified lower bound ($p_L \geq 10.5$) the mean crisis gap is nearly flat (10.9–11.5 pp) and the unemployment skewness varies gently (10.3–11.3), remaining far above HRW's 1.8 throughout. Both soften only when p_L is pushed well below the bound, to 10.2 pp and 7.2 at $p_L = 5$. The vacancy-shutdown mechanism operates throughout (5–6 of the twenty-five grid nodes shut down). The crisis frequency and $\sigma(u)/\sigma(z)$ rise gently with p_L—the latter because the structural parameters are held fixed rather than recalibrated, so the volatility target is met only at the estimate—but neither the sign-reversing gap nor the fat right tail is an artifact of the particular p_L the data fail to pin down from above. The headline results are governed by the *asymmetry* $p_L \gg p_R$, not by the exact left-branch curvature.

Appendix D.3. Sensitivity to the fixed-cost share

The welfare results of Section 6.4 depend on ϕ, the share of total hiring costs attributable to the fixed matching cost κ. The baseline calibration sets $\phi = 0.90$ following Christiano et al. (2016). Because the modified Hosios condition compresses α_L^H toward zero as $\phi \to 1$, the sign-reversing gap documented in Section 6.4 could be sensitive to this choice. To assess robustness, we recalibrate the entire model at $\phi = 0.80$, holding the three simulation targets ($\sigma(u)/\sigma(z)$, $\bar{\theta}$, $\varepsilon_{w,y}$) fixed. The three-moment SMM re-estimates (b, α_L, k): lowering ϕ shifts the efficient bargaining share upward, so the calibration raises α_L to match the same wage elasticity. It also requires a higher b to generate the same unemployment volatility, because less of the hiring cost is sunk and more responds to congestion.

Table D.2: Sensitivity to fixed-cost share ϕ: calibrated parameters and welfare objects.

		$\phi = 0.80$	$\phi = 0.90$
2P-SMCD	b	0.899	0.813
	α_L	0.131	0.112
	Mean Hosios wedge	-0.103	-0.021
	Mean crisis gap (pp)	+10.2	+11.1
	Skewness of u	10.5	10.7
HRW	b	0.848	0.698
	α_L	0.083	0.061
	Mean Hosios wedge	-0.030	+0.001
	Mean crisis gap (pp)	+4.6	+6.7
	Skewness of u	1.8	1.8

Note: At each ϕ, (b, α_L, k) recalibrated to the same three targets. Mean Hosios wedge $= \bar{\alpha}_L - \overline{\alpha_L^H}$; mean crisis gap $= \overline{u_{\text{dec}} - u_{\text{plan}}}$ when $u_{\text{dec}} \geq 15\%$. Skewness: level skewness of the monthly unemployment rate.

Table D.2 reports the results. At $\phi = 0.80$, α_L rises from 0.112 to 0.131 (2P-SMCD) and from 0.061 to 0.083 (HRW), confirming that the low calibrated bargaining weight under $\phi = 0.90$ is driven by the high fixed-cost share rather than an intrinsic feature of the matching function. Matching the same unemployment volatility at lower ϕ requires higher b (0.899 versus 0.813 for the 2P-SMCD; 0.848 versus 0.698 for HRW): with more of the hiring cost responding to congestion, a higher outside option is needed to keep the surplus thin enough for amplification—the Hagedorn and Manovskii (2008) channel, which substitutes for the fixed-cost channel as ϕ falls. The crisis-episode gap is virtually unchanged (+10.2 versus +11.1 pp under the 2P-SMCD) and unemployment skewness remains far higher than HRW's (10.5 versus 1.8). These objects depend on the matching function's *shape*—specifically p_L and θ_c—and are invariant to the cost decomposition. The mean Hosios wedge widens from -0.021 to -0.103 because α_L^H rises with lower ϕ (from 0.134 to 0.234) while α_L rises less, but the qualitative verdict—countercyclical wedge, sign-reversing unemployment gap, fat-tailed crises under the 2P-SMCD—is robust to the choice of fixed-cost share.

Appendix D.4. Solution accuracy

The crisis objects of Section 6—the skewness of unemployment, the crisis frequency, and the crisis-episode gap—are tail statistics whose reliability rests on two things: a simulation long enough that these high-variance statistics have converged, and a global solution that is accurate near the two kinks in the ergodic set (the vacancy-shutdown corner and the crossover θ_c), where linear interpolation is least reliable. We document both.

Convergence of the tail statistics. Level skewness and the crisis objects are dominated by a handful of extreme spells, so at short lengths they are noisy and biased downward: a 200,000-month path under-samples the deep tail. Table D.3 recomputes them on nested prefixes of one long simulation. They converge by about five million months, beyond which the seed is immaterial. This is why Section 6 reports the tail objects from a 5,000,000-month simulation, while the low-variance second moments use the 200,000-month calibration draw. At 200,000 months the 2P-SMCD skewness is an upward outlier (11.6); by five million it settles near 10.7, still an order of magnitude above HRW's 1.8.

Table D.3: Convergence of the tail statistics with simulation length.

	2P-SMCD			HRW	
Months	skew(u)	gap (pp)	P_{crisis} (%)	skew(u)	gap (pp)
200,000	11.64	11.7	0.23	1.78	6.2
1,000,000	9.62	10.7	0.20	1.66	6.3
5,000,000	10.69	11.1	0.22	1.83	6.7
10,000,000	10.68	11.1	0.22	1.85	6.8

Note: Statistics on the leading T months of one 10-million-month simulation (seed 123); successive rows are nested. gap = mean crisis-episode gap ($u_{\text{dec}} - u_{\text{plan}}$ when $u_{\text{dec}} \geq 15\%$). The 5,000,000 row is the reporting length of Section 6.

Euler-equation residuals. At each productivity level z the free-entry condition requires $E(z) = \beta\,\mathbb{E}_{z'|z}[(z'-w')+(1-\delta)E(z')]$. We recompute the right-hand side by the same Gauss–Hermite operator used in the solver, reconstructing the policy from E exactly as the simulation does, and report the unit-free residual $R(z) = \mathrm{RHS}/E(z) - 1$. We use the E-relative residual rather than one expressed in tightness, because $\theta \to 0$ at the shutdown corner, where a tightness-relative residual would diverge, whereas $E(z) \geq \kappa > 0$ stays bounded. Evaluated at the simulated data—which, drawn from the continuous productivity process, fall between grid nodes and so expose interpolation error—the residual averages $10^{-5.8}$ for the 2P-SMCD and $10^{-6.8}$ for HRW. As expected, residuals are largest near the corner where crises occur. Even the worst case, however, stays below 10^{-3} for the 2P-SMCD (an error under 0.1%) and below 10^{-4} for HRW. A fine deterministic grid spanning the state, which guarantees coverage of the deep tail, gives the same maxima. The planner solution, whose bargaining weight $\alpha_L^H(\theta)$ adds curvature of its own, is comparably accurate (worst case $10^{-2.6}$). Holding the shock path fixed at the 200,000-month length, refining the productivity grid from 25 to 200 nodes moves the skewness by less than 0.15 and the crisis gap by under 0.1 pp, so the coarseness of the grid does not drive the results. The mean residual falls monotonically as the grid refines, while its maximum holds near the interpolation error at the kink.

den Haan–Marcet test. Were the solved policy exact, the realized one-step forecast error $e_{t+1} = \beta[(z_{t+1}-w_{t+1})+(1-\delta)E(z_{t+1})]-E(z_t)$ would be orthogonal to any information dated t. We test this by regressing e_{t+1} on powers of the state $\log z_t$; under the null the statistic $T{\cdot}R^2$ is distributed $\chi^2(3)$. Because a single long simulation has near-unbounded power—any nonzero approximation error becomes "significant" as $T \to \infty$—we follow Den Haan and Marcet (1994) and report the rejection frequency across 500 independent samples of 2,500 months. It is 4.4% for the 2P-SMCD and 5.4% for HRW at the nominal 5% level, with a mean statistic near the three degrees of freedom (2.7 and 2.9): the solution's forecast errors are not systematically predictable. As a complementary check on sampling noise, recomputing the crisis objects across 400 independent 200,000-month samples separates the two specifications' unemployment skewness by 5.4 sampling standard deviations, so the gap between them is not a feature of a single draw.