ADDITIONAL SUPPLEMENT TO "EVALUATING POLICY INSTITUTIONS –150 YEARS OF US MONETARY POLICY–"

REGIS BARNICHON
Federal Reserve Bank of San Francisco and CEPR

GEERT MESTERS
Federal Reserve Bank of New York

Abstract We provide additional theoretical results, discuss possible extensions and report additional empirical results and background documents on our monetary policy shock proxy from large gold mine discoveries: S5: Role of news shocks, S6: Functional approximation for "missing path problem", S7: Policy evaluation under constraints, S8: Additional empirical results, S9: Background documents on Gold production by US state.

S5. ROLE OF NEWS SHOCKS

Recall that Lemma 1 in the main text implies

$$\mathbf{Y} = \Gamma(\phi,\theta)\boldsymbol{\Xi} + \mathcal{R}(\phi,\theta)\boldsymbol{\epsilon} \qquad \text{and} \qquad \mathbf{P} = \Gamma_p(\phi,\theta)\boldsymbol{\Xi} + \mathcal{R}_p(\phi,\theta)\boldsymbol{\epsilon}\ , \tag{S1}$$

where $\boldsymbol{\epsilon} = (\epsilon_0', \epsilon_1', \ldots)'$ and $\boldsymbol{\Xi} = (\xi_0', \xi_1', \cdots)'$ are sequences of policy and non-policy shocks, respectively. This shows that in order to identify Γ's and $\mathcal{R}$'s we require knowledge of the current ϵ_0, ξ_0 and future shocks ϵ_h, ξ_h for $h \geq 1$.

Is useful to clarify that in practice this requires the identification of news shocks. To see this, note that we can decompose ξ_t and ϵ_t as[1]

$$\xi_t = \sum_{j=0}^{t} \underbrace{\mathbb{E}_j\xi_t - \mathbb{E}_{j-1}\xi_t}_{\xi_{t,j}} \qquad \text{and} \qquad \epsilon_t = \sum_{j=0}^{t} \underbrace{\mathbb{E}_j\epsilon_t - \mathbb{E}_{j-1}\epsilon_t}_{\epsilon_{t,j}}\ , \tag{S2}$$

where $\mathbb{E}_j(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_j)$, with $\mathcal{F}_j$ the information set available at time j. The increment $\xi_{t,j} \equiv \mathbb{E}_j\xi_t - \mathbb{E}_{j-1}\xi_t$ is the component of ξ_t that is released at time $j \leq t$. In other words $\xi_{t,j}$ is a news shock released at $j \leq t$, and (S2) decomposes the shock ξ_t —a shock realized at time t— as a sum of news shocks $\xi_{t,j}$ revealed all the way until time t with $\xi_t = \sum_{j=0}^t \xi_{t,j}$. Similarly for $\epsilon_{t,j}$. By construction the news shocks are serially uncorrelated.

Thus, to identify the impulse responses in (S1), we require observing proxies for the news shocks in $\boldsymbol{\xi}_0 = (\xi_{0,0}, \xi_{1,0}, \xi_{2,0}, \ldots)'$ and $\boldsymbol{\epsilon}_0 = (\epsilon_{0,0}, \epsilon_{1,0}, \epsilon_{2,0}, \ldots)'$.

For exposition purposes we dropped the zero subscript and worked under perfect foresight.

S6. FUNCTIONAL APPROXIMATIONS FOR "MISSING PATHS"

As discussed in Section 4.3 when we do not identify all policy shocks we cannot evaluate all counterfactual policy paths and our evaluation might be constrained. Possible ways forward include identifying more policy shocks or learning more of $\mathcal{R}^0$ by imposing additional assumptions. In the main text we briefly discussed functional and structural approximations. While the

Regis Barnichon: Regis.Barnichon@sf.frb.org
Geert Mesters: geert.mesters@ny.frb.org

[1] As is common in the optimal policy literature, we impose $\mathbb{E}_{-1}\xi_t = 0$ and $\mathbb{E}_{-1}\epsilon_t = 0$, for all $t = 0, 1, \ldots$. Alternatively, one could let the sums run from $-\infty$ until t.

usage of structural models is discussed in Hebden and Winkler (2021), de Groot et al. (2021) and most notably Caravello et al. (2024), the usage of simple functional approximations has received less attention. For this reason we discuss the advantages and challenges of this approach.

As a starting point, consider the equilibrium outcome for a given variable y_h in Lemma 1, assuming for simplicity that it is scalar. After imposing a unit effect normalization on the scales of the entries of $\boldsymbol{\epsilon}$ we can substitute $\mathbf{P}$ for $\boldsymbol{\epsilon}$ to obtain

$$y_{t+h} = \mathcal{R}_h^0 \mathbf{P}_t + u_{t+h} \ ,$$

where we have appended a time subscript for clarity of exposition. The error term u_{t+h} captures all non-policy shocks and initial conditions. The impulse response $\mathcal{R}_h^0$ captures the horizon h response of an exogenous change to the policy path.

In practice, we only observe a subset of all policy shocks, or proxies therefore, denoted by $\boldsymbol{\epsilon}_{\mathcal{S},t}$. When viewing this as an IV problem, we can say that we are under-identified: we have too few instruments $\boldsymbol{\epsilon}_{\mathcal{S},t}$ relative to the number of endogenous variables in $\mathbf{P}_t$.

To restore point identification we note that in prominent macro models the response $\mathcal{R}_h^0$ is smooth across the different news shocks — some evidence is provided below —. This motivates using functional approximations for approximating $\mathcal{R}_h^0$ and thus reducing the number of policy shocks needed. An early example is found in the work of Inoue and Rossi (2018) who use a factor structure in the spirit of Diebold and Li (2006) to reduce the effective dimension of $\mathcal{R}_h^0$ in the context of monetary policy. Here we discuss a more general basis function approximation that is designed to exploit smoothness across the entries of $\mathcal{R}_h$, i.e. across the news shock horizons.

To set this up, let $\{\phi_j\}$ be a known basis such that $\mathcal{R}_h^0 = \sum_{j=1}^{\infty} \alpha_{hj} \boldsymbol{\phi}_j$. Popular examples include using polynomials, Fourier series, B-splines or Radial basis functions. An approximation for the responses across news shocks is obtained by truncating the number of basis functions. Specifically, given that we can identify, say K, policy news shocks, a K-order approximation becomes

$$y_{t+h} = \boldsymbol{\alpha}_{hK}' \mathbf{P}_{t,K} + u_{t+h} \ ,$$

where $\mathbf{P}_K = (\boldsymbol{\phi}_1 \mathbf{P}_t, \dots, \boldsymbol{\phi}_K \mathbf{P}_t)'$ and $\boldsymbol{\alpha}_{hK}' = (\alpha_{h1}, \dots, \alpha_{hK})'$. Provided that the identified shocks $\boldsymbol{\epsilon}_{\mathcal{S},t}$ are sufficiently correlated with the projected policy paths $\mathbf{P}_{t,K}$ — a testable assumption — we can use IV methods to identify $\boldsymbol{\alpha}_{hK}$ and estimate the vector $\mathcal{R}_h^0$ using the truncated sum $\sum_{j=1}^{K} \alpha_{hj} \boldsymbol{\phi}_j$.

Intuitively, this approach has been used numerous times for smoothing impulse responses, that is smoothing across time horizons h (e.g. Plagborg-Møller, 2016, Barnichon and Brownlees, 2018). The difference here is that we smooth across the horizons of the new shocks. To highlight that in this direction the responses are likely to be smooth we conducted a stylized experiment where we fit prominent structural models to the Post Volcker period and obtain the impulse responses to the different news shocks. Subsequently we fit low dimensional approximations $\mathcal{R}_h^0 \approx \sum_{j=1}^{K} \alpha_{hj} \boldsymbol{\phi}_j$ for different choices of ϕ_j and K. We report the average R^2 across horizons h in table S6 for respectively the interest rate, inflation and output gap response to policy shocks. The models are taken from Caravello et al. (2024) and consist of baseline RANK and HANK models as well a behavioral version of each model: BRANK and BHANK.

We find that a simple linear approximation ($K = 2$) does not work well for capturing all responses; notably the response for output is poorly approximated. When we consider a quadratic approximation ($K = 3$) the results improve substantially and when we consider $K = 3$ Gaussian basis functions the fit is nearly perfect for all models considered. This implies that, regardless of the specific model considered, if we could identify three monetary policy shocks that

TABLE S5

FUNCTIONAL APPROXIMATIONS TO MODEL IMPLIED NEW SHOCK RESPONSES

Lin ($\|\mathcal{S}\| = 2$)	rank	brank	hank	bhank
i	0.96	0.90	0.80	0.94
π	0.94	0.69	0.32	0.94
y	0.15	0.10	0.83	0.44
Quad ($\|\mathcal{S}\| = 3$)				
i	1.00	1.00	0.99	1.00
π	0.99	0.96	0.82	0.99
y	0.88	0.72	0.84	0.92
Gauss ($\|\mathcal{S}\| = 3$)				
i	0.99	0.98	0.97	0.98
π	0.99	0.99	0.99	0.97
y	0.97	0.99	0.88	0.98

Note: The table reports the average R^2's obtained from fitting responses $\mathcal{R}_h^0$ obtained from the structural models listed in the columns using basis function approximations.

capture different news horizons we could approximate $\mathcal{R}^0$ using functional approximations. While identifying three —sufficiently distinct— policy shocks remains challenging for today's frontier of empirical macro, this simulation exercise shows that the identification of different policy (or non-policy) shocks is an important direction for improvement in future work.

S7. OPTIMAL REACTION ADJUSTMENT UNDER CONSTRAINTS

In this section, we show how to extend our approach to handle constraints on the policy maker's problem. The policy problem that we considered in the main text is an unconstrained optimization problem: there are no restrictions on the policy path. In practice however, policy makers may face additional constraints, e.g., a lower-bound on the policy rate in the context of monetary policy.

In the main text, we saw that the timeless (unconstrained) minimization of the expected loss is equivalent to the minimization of the (weighted) sum-of-squares of the impulse responses to shocks. However, since impulse responses are defined as differences between two forecasts, we cannot use that approach to impose level restrictions on the policy problems. For instance, from the point of view of an impulse response the "level" of the policy instrument is irrelevant (in a linear framework).

Fortunately, we will now see that there exists an equivalent representation of the unconditional policy problem that allows to incorporate constraints on the policy problem. The intuition is simple. The timeless policy problem can also be seen as the minimization of the unconditional expectation of a time t loss function —$\min_{\phi} \mathbb{E}\mathcal{L}_t$ where $\mathcal{L}_t = \mathbb{E}_t \mathbf{Y}_t^{\prime} \mathcal{W} \mathbf{Y}_t$—, so that we can also find the optimal policy rule by replacing the expectation operator with its finite sample counterpart, that is find the optimal rule that minimizes the average loss over a sequence of repeated decisions: $\min_{\phi} \sum_{t=t_i}^{t_f} \mathcal{L}_t$. Asymptotically, this approach will also deliver ϕ^{opt}. In other words, by solving the time t problem for a *sequence* of policy decisions, we can equivalently characterize the optimal reaction function. This equivalence will then allow us to easily incorporate constraints on the policy instruments that may bind depending on the state of the economy at time t.

S7.1. *The time-t policy problem*

To study the time t problem, we need a slight modification of our environment to explicitly capture initial conditions. A generic model for the non-policy block of the economy at time t is given by

$$\mathcal{A}_{yy}\mathbb{E}_t\mathbf{Y}_t - \mathcal{A}_{yp}\mathbb{E}_t\mathbf{P}_t = \mathcal{B}_{yx}\mathbf{X}_{-t} + \mathcal{B}_{y\xi}\mathbf{\Xi}_t \ , \tag{S3}$$

where the vector $\mathbf{X}_{-t} = (y'_{t-1}, p'_{t-1}, y'_{t-2}, \ldots)'$ captures the initial conditions as summarized by the history of the variables y_t,and p_t, and $\mathbf{\Xi}_t = (\xi'_t, \xi'_{t,t+1}, \xi'_{t,t+2}, \ldots)'$ denotes the path of the structural shocks to the economy. Specifically, ξ_t is the time t vector of structural (non-policy) shocks, while the shocks $\xi_{t,t+h}$, for $h = 1, 2, \ldots$, are news shocks: information revealed at time t about shocks that realize at time $t+h$. The vector $\mathbf{\Xi}_t$ thus includes all shocks that are released at time t. We normalize the news shocks to be mutually uncorrelated with mean zero and unit variance. The linear maps $\mathcal{A}_{..}$ and $\mathcal{B}_{..}$ are conformable and we define the time t information set in terms of the pre-determined inputs as $\mathcal{F}_t = \{\mathbf{X}_{-t}, \mathbf{\Xi}_t\}$.

Similarly, we consider a generic model for the policy block with

$$\mathcal{A}_{pp}\mathbb{E}_t\mathbf{P}_t - \mathcal{A}_{py}\mathbb{E}_t\mathbf{Y}_t = \mathcal{B}_{px}\mathbf{X}_{-t} + \mathcal{B}_{p\xi}\mathbf{\Xi}_t + \boldsymbol{\epsilon}_t \ , \tag{S4}$$

where $\boldsymbol{\epsilon}_t = (\epsilon'_t, \epsilon'_{t,t+1}, \epsilon'_{t,t+2}, \ldots)'$ is the path of policy news shocks. Specifically, the vector $\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_t)'$ includes the contemporaneous policy shocks to the policy instrument and $\epsilon_{t,t+h}$ are policy news shocks: information revealed at time t about policy shocks that realize at time $t+h$. The policy news shocks $\boldsymbol{\epsilon}_t$ are mean zero with unit variance and uncorrelated with the initial conditions $\mathbf{X}_{-t}$ and the other non-policy shocks $\mathbf{\Xi}_t$.

We collect all parameters of the policy rule (S4) in $\phi = \{\mathcal{A}_{pp}, \mathcal{A}_{py}, \mathcal{B}_{px}, \mathcal{B}_{p\xi}\}$. A policy choice is then defined by a pair $(\phi, \boldsymbol{\epsilon}_t)$ consisting of the rule parameters ϕ and the policy news shocks $\boldsymbol{\epsilon}_t$.

Consider some baseline policy choice $(\phi^0, \boldsymbol{\epsilon}_t^0)$ and denote by $\mathbb{E}_t\mathbf{P}_t^0$ and $\mathbb{E}_t\mathbf{Y}_t^0$ the associated baseline paths for the policy instruments and policy objectives. For now, that baseline choice is arbitrary except for the following assumption

Under rule ϕ^0 enduring a unique equilbrium, we have (see Barnichon and Mesters, 2023)

$$\begin{aligned} \mathbb{E}_t\mathbf{Y}_t^0 &= \Gamma_y^0\mathbf{S}_t + \mathcal{R}_y^0\boldsymbol{\epsilon}_t^0 \\ \mathbb{E}_t\mathbf{P}_t^0 &= \Gamma_p^0\mathbf{S}_t + \mathcal{R}_p^0\boldsymbol{\epsilon}_t^0 \end{aligned} \ , \tag{S5}$$

with $\mathbf{S}_t = (\mathbf{X}'_{-t}, \mathbf{\Xi}'_t)'$ and $\mathbb{E}(\boldsymbol{\epsilon}_t^0\mathbf{S}'_t) = 0$. Expression (S5) characterizes the model solution, expressing the endogenous variables $\mathbb{E}_t\mathbf{Y}_t^0$ and $\mathbb{E}_t\mathbf{P}_t^0$ as functions of the state of the economy $\mathbf{S}_t = (\mathbf{X}'_{-t}, \mathbf{\Xi}'_t)'$, which captures initial conditions $\mathbf{X}_{-t}$ and the non-policy shocks $\mathbf{\Xi}_t$, and the policy shocks $\boldsymbol{\epsilon}_t^0$. Note that $\mathbb{E}_t\mathbf{Y}_t^0$ and $\mathbb{E}_t\mathbf{P}_t^0$ are the oracle forecasts as of time t, that is the expectations for $\mathbf{Y}_t$ and $\mathbf{P}_t$ conditional on the information set $\mathcal{F}_t$ and the policy choices $(\phi^0, \boldsymbol{\epsilon}_t^0)$.

The time t perturbation approach consists in starting form some baseline policy choice $(\phi^0, \boldsymbol{\epsilon}_t^0)$ and modifying that baseline policy rule ϕ^0 as follows

$$\mathcal{A}_{pp}^0\mathbb{E}_t\mathbf{P}_t - \mathcal{A}_{py}^0\mathbb{E}_t\mathbf{Y}_t = \mathcal{B}_{px}^0\mathbf{X}_{-t} + \mathcal{B}_{p\xi}^0\mathbf{\Xi}_t + \boldsymbol{\epsilon}_t^0 + \boldsymbol{\delta}_t,$$

where $\boldsymbol{\delta}_t = (\delta_{t,t}, \delta_{t,t+1}, \ldots)'$ is a path of adjustments to the policy rule equations. Given the generic model (14) and the modified policy rule (S7.1), under $(\phi^0, \boldsymbol{\epsilon}_t^0)$ with ϕ^0 ensuring a unique equilibrium, we then have (see Barnichon and Mesters, 2023)

$$\begin{aligned} \mathbb{E}_t\mathbf{Y}_t(\boldsymbol{\delta}_t) &= \mathbb{E}_t\mathbf{Y}_t^0 + \mathcal{R}_y^0\boldsymbol{\delta}_t \\ \mathbb{E}_t\mathbf{P}_t(\boldsymbol{\delta}_t) &= \mathbb{E}_t\mathbf{P}_t^0 + \mathcal{R}_p^0\boldsymbol{\delta}_t \end{aligned} \ . \tag{S6}$$

The optimal policy can be characterized by considering a planner who chooses the paths $\mathbf{Y}_t, \mathbf{W}_t$ and $\mathbf{P}_t$ in order to minimize the loss function, i.e.,

$$\min_{\mathbf{Y}_t, \mathbf{P}_t} \mathcal{L}_t \qquad \text{s.t.} \qquad (14)\,. \tag{S7}$$

Denote by $\mathbb{E}_t \mathbf{P}_t^{\text{opt}}$ the corresponding optimal policy path. Note that the problem defines the entire optimal policy path as a function of the information available at time t.

To find the optimal policy path, we can exploit the "law of motion" (S6) to search for the "best" adjustment $\boldsymbol{\delta}_t$ to the baseline rule ϕ^0, that is find the $\boldsymbol{\delta}_t$ that solves the problem

$$\boldsymbol{\delta}_t^* = \underset{\boldsymbol{\delta}_t}{\operatorname{argmin}}\ \mathcal{L}_t(\boldsymbol{\delta}_t) \qquad \text{s.t.} \qquad \mathbb{E}_t \mathbf{Y}_t(\boldsymbol{\delta}_t) = \mathbb{E}_t \mathbf{Y}_t^0 + \mathcal{R}_y^0 \boldsymbol{\delta}_t\,,$$

where $\mathcal{L}_t(\boldsymbol{\delta}_t) = \mathbb{E}_t \mathbf{Y}_t(\boldsymbol{\delta}_t)' \mathcal{W} \mathbb{E}_t \mathbf{Y}_t(\boldsymbol{\delta}_t)$. This perturbed policy problem has a closed form solution given by

$$\boldsymbol{\delta}_t^* = -(\mathcal{R}_y^{0'} \mathcal{W} \mathcal{R}_y^0)^{-1} \mathcal{R}_y^{0'} \mathcal{W} \mathbb{E}_t \mathbf{Y}_t^0\,. \tag{S8}$$

The perturbation $\boldsymbol{\delta}_t^*$ is the *Optimal Policy Perturbation* (OPP) introduced in Barnichon and Mesters (2023). It allows to characterize the optimal policy path for the time t policy problem, that is $\mathbb{E}_t \mathbf{P}_t^{\text{opt}}$.

From OPP to ORA

The OPP is useful in our present context (a timeless perspective on policy evaluation) because combining (S5) and (S8) shows a link between OPP and ORA. Specifically, we have

$$\boldsymbol{\delta}_t^* = -(\mathcal{R}_y^{0'} \mathcal{W} \mathcal{R}_y^0)^{-1} \mathcal{R}_y^{0'} \mathcal{W} \left(\Gamma_y^0 \mathbf{S}_t + \mathcal{R}_y^0 \boldsymbol{\epsilon}_t^0 \right) \text{ where } \mathbf{S}_t = (\mathbf{X}_{-t}', \boldsymbol{\Xi}_t')'. \tag{S9}$$

Taking ξ_t some arbitrary element of $\boldsymbol{\Xi}_t$, we then have

$$\begin{aligned} \boldsymbol{\delta}_t^* &= -(\mathcal{R}_y^{0'} \mathcal{W} \mathcal{R}_y^0)^{-1} \mathcal{R}_y^{0'} \mathcal{W} \Gamma_{y,\xi}^0 \xi_t + \text{other terms independent of } \xi_t \\ &= \boldsymbol{\tau}_\xi^* \xi_t + \text{other terms independent of } \xi \end{aligned} \tag{S10}$$

where τ_ξ is the ORA for non-policy shock ξ.

Expression (S10) shows that it is possible to estimate the ORA from (i) a sequence of OPP $\{\boldsymbol{\delta}_t^*\}_{t=t_i}^{t_f}$ and a sequence of the non-policy shock ξ_t, i.e., $\{\xi_t\}_{t=t_i}^{t_f}$. Specifically, we can estimate $\boldsymbol{\tau}_\xi$ consistently from the regressions:[2]

$$\boldsymbol{\delta}_t^* = \gamma \xi_t + \boldsymbol{u}_t. \tag{S11}$$

While we recommend estimating the ORA from expression (23) in the main text (a simple impulse response regression), we will now see how regression (S11) can be easily adapted to estimate the ORA (and then the DML and optimal policy response path adjustment) in the case of *constrained* optimization problems, when the constraint depends on the nature of the time t problem. We first describe the constrained optimization problem and then show how to estimate the corresponding *constrained* ORA.

[2]The estimator is consistent because the residual is independent of ξ_t.

S7.2. *The constrained time-t policy problem*

We allow the constraints to be general nonlinear functions of $\mathbb{E}_t\mathbf{Y}_t$ and $\mathbb{E}_t\mathbf{P}_t$ that can be written as

$$C(\mathbb{E}_t\mathbf{Y}_t, \mathbb{E}_t\mathbf{P}_t) \geq \mathbf{c}\ , \tag{S12}$$

where $C(\cdot,\cdot) : \mathbb{R}^\infty \times \mathbb{R}^\infty \to \mathbf{R}^{d_c}$ is the known constraint function and $\mathbf{c}$ a vector of constants of length d_c (possibly infinite). As an example, suppose that $\mathbb{E}_t\mathbf{P}_t = \mathbb{E}_t(i_t, i_{t+1}, \ldots)$ is the expected interest rate path of a central bank. We can impose the zero lower bound by setting $C(\mathbb{E}_t\mathbf{Y}_t, \mathbb{E}_t\mathbf{P}_t) = \mathbb{E}_t\mathbf{P}_t$ and $\mathbf{c} = \mathbf{0}$.

To incorporate the constraints into the policy problem we modify the time-t policy problem (S7) to become

$$\min_{\mathbf{Y}_t,\mathbf{P}_t} \mathcal{L}_t \qquad \text{s.t.} \qquad (14) \qquad \text{and} \qquad C(\,\mathbb{E}_t\mathbf{Y}_t, \mathbb{E}_t\mathbf{P}_t) \geq \mathbf{c}\ . \tag{S13}$$

The optimal solution for $\mathbb{E}_t\mathbf{P}_t$ is denoted by $\mathbb{E}_t\mathbf{P}_t^{\text{opt},c}$, and for simplicity we assume that it is unique. Following the same steps as with the unconstrained OPP, we can construct a *constrained OPP* statistic given by

$$\begin{aligned} \boldsymbol{\delta}_t^{c*} = \underset{\boldsymbol{\delta}_t}{\text{argmin}}\ \mathcal{L}_t(\boldsymbol{\delta}_t) \qquad & \text{s.t.} \quad \mathbb{E}_t\mathbf{Y}_t(\boldsymbol{\delta}_t) = \mathbb{E}_t\mathbf{Y}_t^0 + \mathcal{R}_y^0\boldsymbol{\delta}_t\ , \\ & \text{and} \quad C(\mathbb{E}_t\ \ \mathbf{Y}_t^0 + \mathcal{R}_y^0\boldsymbol{\delta}_t, \mathbb{E}_t\mathbf{P}_t^0 + \mathcal{R}_p^0\boldsymbol{\delta}_t) \geq \mathbf{c}\ . \end{aligned} \tag{S14}$$

In contrast to the baseline OPP statistic (S8), there exists no closed form solution for $\boldsymbol{\delta}_t^{c*}$. Nevertheless, we can easily solve this problem numerically as all inputs are the same as above.

From constrained OPP to constrained ORA

Proceeding as in (S10), we can estimate the constrained ORA for some non-policy shock ξ_t from the regressions

$$\boldsymbol{\delta}_t^{c*} = \boldsymbol{\gamma}\xi_t + \boldsymbol{u}_t, \tag{S15}$$

that is, we can estimate $\boldsymbol{\tau}_\xi^{c*}$ from a sequence of constrained OPP $\{\boldsymbol{\delta}_t^{c*}\}_{t=t_i}^{t_f}$ and a sequence of the non-policy shock ξ_t, i.e., $\{\xi_t\}_{t=t_i}^{t_f}$. From the constrained ORA, we can then construct the constrained optimal policy response path adjustment from $\Delta^c\Theta_{p,s} = -\mathcal{R}_p^0\boldsymbol{\tau}_s^{c*}$ and the constrained DML $\Delta^c\mathcal{L} = \left(\mathcal{R}^0\boldsymbol{\tau}^{c*}\right)' \mathcal{W}\mathcal{R}^0\boldsymbol{\tau}^{c*}$.

Note that sufficient statistics requirements are the exact same as in the unconstrained ORA case, since estimating an impulse response to non-policy shock (Γ_ξ) requires the identification of a sequence of non-policy shock ξ_t.

S8. ADDITIONAL ROBUSTNESS CHECKS

In this section, we report the results of a number of robustness checks: (i) evaluating the Post Volcker regime with a zero lower bound constraint on the policy rate, (ii) robustness to time-variation in the natural rate of unemployment (u^*), (iii) an alternative identification of monetary shocks (set identification based on sign restrictions). In the interest of space, we will report the robustness results for the ORA estimates only, as the ORA underlies the distance to minimum loss and optimal adjustment to the policy path response.

S8.1. *Evaluating the Post Volcker regime under a ZLB constraint*

During the 2007-2008 crisis, the Fed's response was arguably constrained by the zero lower bound as the fed funds rate remained at the ZLB from December 2007 until December 2015. To take this constraint into account when evaluating policy performance, Table S6 reports our estimates for the ZLB-constrained ORA following (S15).[3]

The results are similar to our baseline results —no ORA is significantly different from zero—, but note how the ORA to financial shocks is no longer negative. In the unconstrained case, the ORA point estimate was negative at -0.3, which indicated too little monetary stimulus in the face of adverse financial shock (though the result was not statistically significant). Table S6 shows that this result disappears once we incorporate the zero-lower bound constraint. In other words, this suggests that the ZLB did constrain somewhat the Fed's reaction function during the 2007-2008 financial crisis, though the deviation from the optimal unconstrained response is small.

TABLE S6
CONSTRAINED ORA STATISTICS

Non-policy shock Shock sign convention	Bank panics $u\uparrow$	G $u\uparrow$	Energy $\pi\uparrow$	π^e $\pi\uparrow$	TFP $\pi\uparrow$
Post Volcker 1990–2019	**0.1** $(-0.2,0.3)$	**−0.3** $(-4.6,4.0)$	**−0.1** $(-0.6,0.5)$	**0.0** $(-0.1,0.1)$	**−0.0** $(-0.1,0.1)$

Note: Median ORA statistics together with 68% credible sets. See the main text for shock identification assumptions.

S8.2. *Time-variation in u^**

We check the robustness of our results to time-variation in the natural rate of unemployment. Instead of using a constant u^* as in the main text, we used time-varying estimates of u^*, either (Barnichon and Matthes, 2017)'s estimate from a TVP-VAR (which effectively models u_t^* as following a random walk) or (Ramey and Zubairy, 2018)'s estimate (which draws on the CBO's estimate post World War II and on a low-frequency filter pre World War II). While there exists no universal u_t^* measures, Table S7 shows that we obtain very similar results across the different u_t^* estimates, almost identical to our main text baseline results (which treat the policy makers' target as a constant).

S8.3. *Category-specific comparison: policy response to forecast revisions*

In this section, we consider an alternative categorical comparison based on forecast revisions, i.e., assessing how well policy makers responded to Wold innovations to the BAA-AAA spread, government spending, energy inflation, inflation expectations and labor productivity growth. To that effect, we use Corollary 3 from the main text to compute the Wold-based subset DML

[3]The post Volcker period is the only period where the ZLB constraint was actually binding, even after ORA adjustment.

TABLE S7

ORA STATISTICS, ROBUSTNESS TO TIME-VARYING u_t^*

	Bank panics	G	Energy	π^e	TFP
		u_t^* from Ramey and Zubairy (2018)			
Pre Fed 1879−1912	**−0.7*** (−1.4,−0.1)	**0.5** (−0.2,1)	**0.0** (−0.4,0.5)	—	—
Early Fed 1913−1941	**−1.2*** (−2,−0.8)	**−0.4*** (−0.9,−0.1)	**0.0** (−0.4,0.3)	**0.7*** (0.3,1.1)	—
Post WWII 1951−1984	—	**−0.2** (−0.8,0.4)	**0.7*** (0.1,1.3)	**1.2*** (0.5,2)	**0.5** (−0.2,1.1)
Post Volcker 1990−2019	**0.2** (−0.5,0.9)	**0.3** (−0.2,0.8)	**0.0** (−0.8,0.9)	**0.0** (−0.5,0.4)	**−0.2** (−0.6,0.2)
		u_t^* from Barnichon and Matthes (2017)			
Early Fed 1913−1941	**−1.4*** (−2.2,−0.9)	**−0.5*** (−1,−0.1)	**0.0** (−0.4,0.3)	**0.7*** (0.3,1.1)	—
Post WWII 1951−1984	—	**−0.2** (−0.8,0.4)	**0.6*** (0.1,1.2)	**1.3*** (0.6,2.1)	**0.5** (−0.1,1.2)
Post Volcker 1990−2019	**−0.1** (−0.6,0.5)	**0** (−0.6,0.6)	**0.1** (−0.6,0.7)	**−0.1** (−0.5,0.3)	**−0.3** (−0.7,0.2)

Note: Median ORA statistics together with 68% credible sets. See the main text for shock identification assumptions. u_t^* is measured from a TVP-VAR as in Ramey and Zubairy (2018) (top panel) or from Barnichon and Matthes (2017) (bottom panel).

$\Delta_S \mathcal{L}_{U_c}$ from a large VAR, similarly to the one used in Section 5: a VAR with four lags and seven variables: inflation, unemployment, the identified policy shock, energy inflation, labor productivity growth, the BAA-AAA spread, and government spending (as share of potential output).

TABLE S8

CATEGORY-SPECIFIC DML: POLICY RESPONSE TO FORECAST REVISIONS

	Subset DML ($\Delta_S \mathcal{L}_{U_c}$)				
	BAA-AAA	G	Energy	π^e	TFP
Pre Fed 1879−1912	**0.8** (0.1,5.3)	**0.5** (0.1,3.7)	—	—	—
Early Fed 1913−1941	**16.8** (4.7,57.6)	**7.6** (0.6,65.7)	**1.7** (0.2,9.1)	**4.4** (0.5,20)	—
Post WWII 1951−1984	**0.2** (0,1.0)	**0.1** (0,0.5)	**0.3** (0,1.4)	**1.6** (0.3,5)	**0.1** (0,0.7)
Post Volcker 1990−2019	**0.2** (0,1.1)	**0.1** (0,0.4)	**0.1** (0,0.8)	**0.2** (0,1.2)	**0.1** (0,0.9)

Note: Median subset distance to minimum loss together with 68% credible sets in parentheses with each row reporting estimates for a different period.

Table S8 shows the results. The policy performance comparisons are in line with our main conclusions in the main text: substantial improvement in the reaction to aggregate demand type innovations (financial or government spending) after World War II, and substantial improvement in the reaction to inflation expectation or energy price innovations after Volcker. We can also confirm how the Early Fed fared worse than the passive gold standard in the face of financial and government spending innovations.

S8.4. *Alternative monetary shocks identifying assumptions*

We now consider robustness to the identification of monetary shocks, and table shows the ORA statistics estimated using sign-identified monetary policy shocks. The results are remarkably consistent with our baseline estimates, with ORAs generally of similar magnitudes and same levels of statistical significance.

TABLE S9

ORA STATISTICS, SIGN-BASED IDENTIFICATION

Non-policy shock Shock sign convention	Bank panics $u\uparrow$	G $u\uparrow$	Energy $\pi\uparrow$	π^e $\pi\uparrow$	TFP $\pi\uparrow$
Pre Fed 1879−1912	**−0.6*** (−0.9,−0.3)	**−0.4*** (−0.7,−0.2)	**0.0** (−0.4,0.4)	—	—
Early Fed 1913−1941	**−0.8*** (−1.2,−0.4)	**−0.4*** (−0.7,0)	**0.0** (−0.3,0.3)	**0.7*** (0.4,1.1)	—
Post WWII 1951−1984	—	**−0.3** (−0.6,0.1)	**0.5** (−0.1,1.1)	**0.7*** (0.3,1.1)	**0.5*** (0.0,1.0)
Post Volcker 1990−2019	**−0.3*** (−0.7,0.0)	**0.3** (−0.2,0.8)	**−0.3** (−0.8,0.3)	**0.2** (−0.2,0.5)	**0.2** (−0.1,0.5)

Note: Median ORA statistics together with 68% credible sets. See the main text for shock identification assumptions.

S9. BACKGROUND DOCUMENTS ON GOLD PRODUCTION BY US STATE

To inform our narrative identification of large gold mine discoveries and peak extraction, we rely on Koschmann and Bergendahl (1968), which is a detailed account of gold production districts in the US since 1799. Figures S22-S25 plot historical gold production at the state level for the major Gold producers over 1870-1913. Gold rushes led to large variations in gold production; in the order of 30-40 percent of *national* production.

BIBLIOGRAPHY

BARNICHON, REGIS AND CHRISTIAN BROWNLEES (2018): "Impulse Response Estimation By Smooth Local Projections," *Review of Economics and Statistics*. [2]

BARNICHON, REGIS AND CHRISTIAN MATTHES (2017): "The natural rate of unemployment over the past 100 years," *FRBSF Economic Letter*, 23, 219–231. [7, 8]

BARNICHON, REGIS AND GEERT MESTERS (2023): "A Sufficient Statistics Approach for Macroeconomic Policy," *American Economic Review*, 113 (11), 2809–45. [4, 5]

CARAVELLO, TOM, ALISDAIR MCKAY, AND CHRISTIAN WOLF (2024): "Evaluating Policy Counterfactuals: A "VAR-Plus" Approach," . [2]

DE GROOT, OLIVER, FALK MAZELIS, ROBERTO MOTTO, AND ANNUKKA RISTINIEM (2021): "A toolkit for computing Constrained Optimal Policy Projections (COPPs)," ECB working paper: No 2555. [2]

DIEBOLD, FRANCIS X. AND CANLIN LI (2006): "Forecasting the term structure of government bond yields," *Journal of Econometrics*, 130 (2), 337–364. [2]

HEBDEN, JAMES AND FABIAN WINKLER (2021): "Impulse-Based Computation of Policy Counterfactuals," Finance and Economics Discussion Series 2021-042. Washington: Board of Governors of the Federal Reserve System. [2]

INOUE, ATSUSHI AND BARBARA ROSSI (2018): "The Effects of Conventional and Unconventional Monetary Policy: A New Approach," Working paper. [2]

KOSCHMANN, ALBERT HEINRICH AND MAXIMILIAN H BERGENDAHL (1968): "Principal Gold-Producing Districts of the United States: a description of the geology, mining history, and production of the major gold-mining districts in 21 States," *Professional paper/US Department of the Interior, US Geological Survey*. [9, 11, 12]

PLAGBORG-MØLLER, M. (2016): "Essays in Macroeconometrics, Chapter 3," PhD Dissertation, Harvard University. [2]

RAMEY, VALERIE A. AND SARAH ZUBAIRY (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data," *Journal of Political Economy*, 126. [7, 8]

FIGURE S22.— Alaska Gold Production

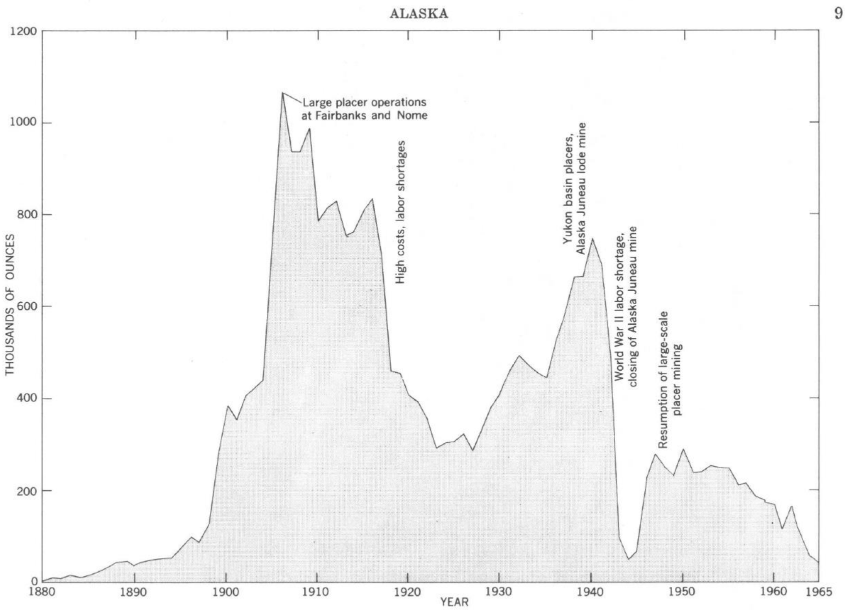
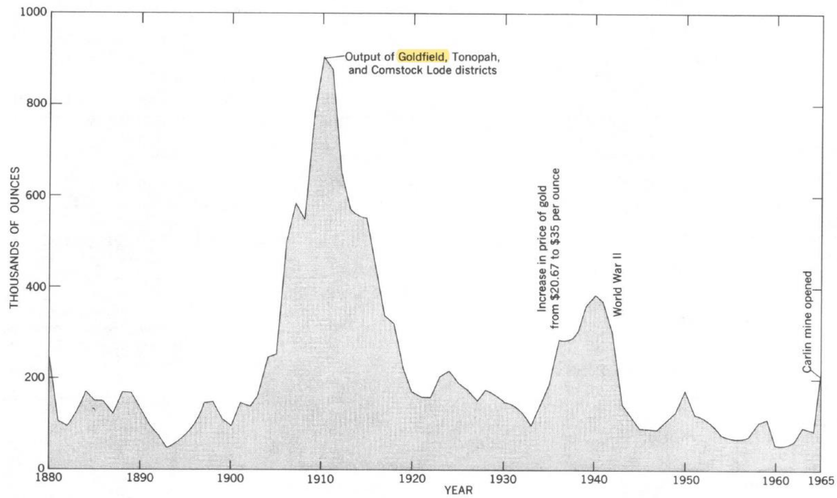


FIGURE 4.—Annual gold production of Alaska, 1880–1965. Sources of data: 1880–1900, U.S. Geological Survey (1883–1924); 1900–42, Smith (1944, p. 6); 1943–59, U.S. Bureau of Mines (1933–66). Production reported in dollar value was converted to ounces at prevailing price per ounce.

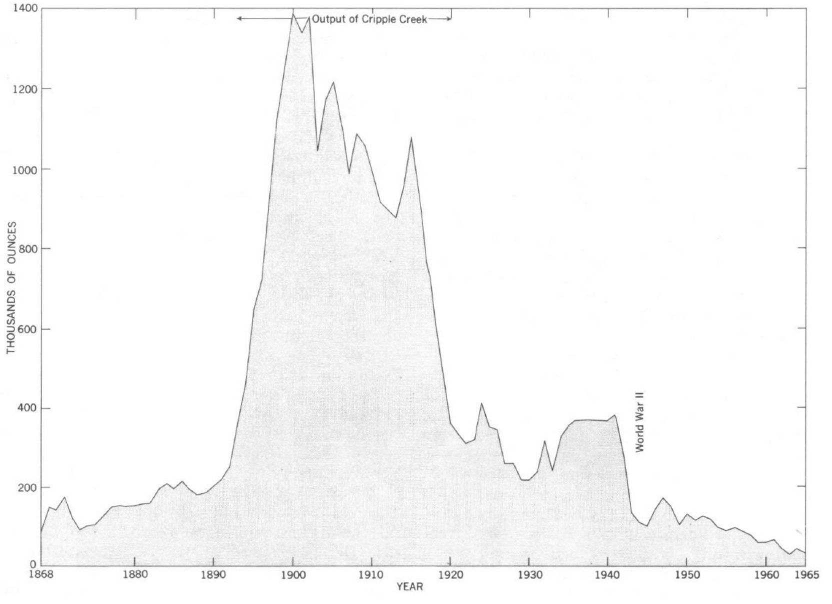
Note: *Source*: Koschmann and Bergendahl (1968).

FIGURE S23.— Nevada Gold Production



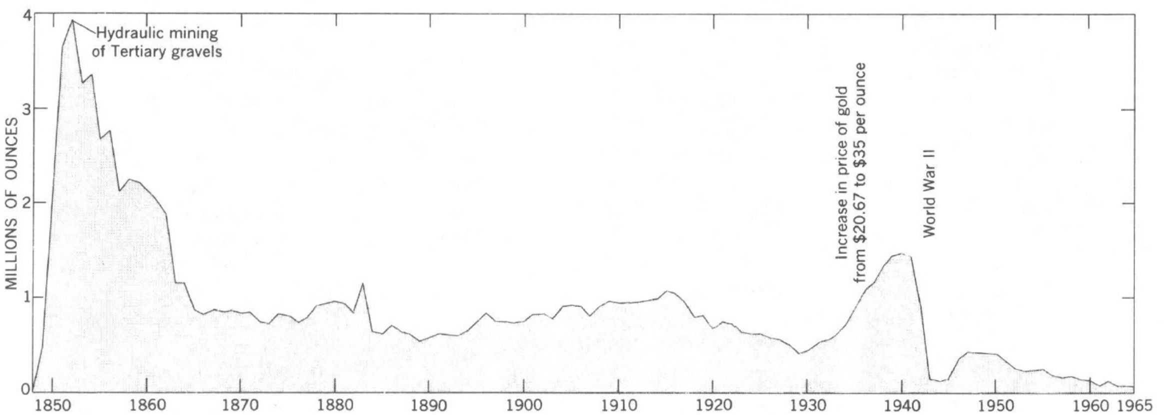
Note: *Source*: Koschmann and Bergendahl (1968).

FIGURE S24.— Colorado Gold Production



Note: *Source*: Koschmann and Bergendahl (1968).

FIGURE S25.— California Gold Production



Note: *Source*: Koschmann and Bergendahl (1968).