OUTPUT GAP ESTIMATION AND MONETARY POLICY WITH IMPERFECT KNOWLEDGE*

Pei Kuang† Kaushik Mitra‡ Li Tang§

October 2026

Abstract

The output gap (OG) is a key input into monetary policy, but being unobserved, alternative real-time estimation methods are used to provide its estimates. We compare eleven widely used methods in a New Keynesian model with adaptive learning, in which the policymaker's real-time OG estimates feed back into interest rates, aggregate outcomes, and subsequent estimates. The choice of method matters substantially for macroeconomic stability and welfare, particularly when the zero lower bound binds: Beveridge–Nelson-type and bandpass filters remain stable across a wide range of policy rules, whereas deterministic detrending invites self-reinforcing deflation spirals.

JEL classification: C18, E17, E32, E52

Keywords: Detrending, Monetary policy, Expectations, Learning, Inflation, Welfare

*We thank for helpful discussions and comments Klaus Adam, James Bullard, Yuriy Gorodnichenko, Gary Koop, Leonardo Melosi, Nigel McClung, Athanasios Orphanides, Sergey Slobodyan, Vincent Sterk, Simon van Norden, Fabian Winkler, seminar participants at the University of Osaka, University of Illinois Urbana-Champaign, and participants of the 2021 Conference on Expectations in Dynamic Macroeconomic Models (Czech National Bank), 2022 ESCoE Conference on Economic Measurement (University of Strathclyde, Glasgow), 2022 Central Bank Modelling workshop (Banco de la República), and 2022 South-Eastern European Economic Research Workshop (Bank of Albania). The computations described in this paper were performed using the University of Birmingham's supercomputer (BlueBEAR) high performance computing service.

†Department of Economics, University of Macau, E21B Avenida da Universidade, Taipa, Macau. E-mail: peikuang@um.edu.mo

‡Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom. E-mail: K.Mitra@bham.ac.uk

§Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom. E-mail: LiTang.Econ@gmail.com

1 Introduction

Central banks around the world devote significant resources to estimating the output gap (OG) in real time, as it is a key input into monetary policy decisions.[1] The OG, however, is not directly observed, and the many methods developed for its estimation—univariate filters, multivariate decompositions, structural VARs—often deliver substantially different estimates from the same data. Macroeconomic policy analysis nevertheless proceeds largely as if the choice among them were innocuous: models are analyzed in detrended form, with limited discussion of the detrending method behind the detrended equations. This paper takes that choice seriously as a primitive of monetary policy design.

We embed eleven commonly recommended OG estimation methods in a New Keynesian (NK) model with learning, in which the policymaker estimates the OG in real time and responds to it. Each method is thereby evaluated inside the dynamic feedback loop it creates—OG estimates shape interest rates and aggregate outcomes, which in turn shape subsequent estimates—rather than against exogenously given data, and we characterize the welfare and monetary policy implications of each method. Our central result is that the choice of estimation method matters substantially for macroeconomic stability and welfare—and matters most precisely when policy is most constrained, at the zero lower bound (ZLB): Beveridge–Nelson-type filters, the bandpass filter, the Hodrick–Prescott (HP) filter and first-order differencing keep the economy stable over a wide range of interest rate rules, whereas deterministic detrending methods (linear, quadratic and discounted least squares) expose it to self-reinforcing deflation spirals and large welfare losses over much of the policy space.

The model is an NK economy with self-referential learning, as in Marcet and Sargent (1989) and Eusepi and Preston (2018). The central bank commits to a simple Taylor-type rule (see e.g. Woodford, 2003) in which the interest rate responds to the deviation of inflation from target and to the real-time OG estimate (and possibly the lagged interest rate), subject to the ZLB. The modelling innovation is that the policymaker forms this real-time estimate with one of the eleven estimation methods. Monetary policy is evaluated with a standard welfare loss function, a weighted average of the unconditional variances of inflation and the OG.

Our headline criterion for comparing methods is the size of the "stability region": the share

[1] Interest in OG estimation was renewed after the global financial crisis, e.g., by the secular stagnation debate (Summers, 2014) and by the role of cyclically adjusted budget balances—in practice estimated using OG measures—in EU fiscal surveillance (Kuang and Mitra, 2025). While OG estimates are crucial for both monetary and fiscal policy, this paper focuses on the monetary policy channel.

of interest rule coefficients for which the associated welfare loss remains below a threshold, so that the economy avoids self-reinforcing inflation or deflation spirals. A large stability region is desirable because neither the policy coefficients a central bank adopts in practice nor the economy's structural parameters are known precisely; a method that remains stable over a wide region guards against policy or estimation error tipping the economy into instability.[2]

Our first and most important finding is that the methods are far from interchangeable, and their differences follow a clear pattern. With the ZLB and no interest rate inertia, the bandpass filter of Christiano and Fitzgerald (2003), the Beveridge–Nelson filter with a low signal-to-noise ratio of Kamber et al. (2018) (BN-KMW), the one-sided HP filter and first-order differencing (FOD) generate the largest stability regions—between 59% and 80% of the policy parameter space—whereas the deterministic methods perform worst: quadratic and linear detrending are stable on less than a quarter of the parameter space, and discounted least squares (DLS) on a mere 2%. Methods prone to instability share a common trait: they intrinsically extract large and volatile OG estimates,[3] which feed into interest rates, amplify the volatility of inflation and beliefs, and raise the likelihood that the ZLB binds and deflation spirals take hold. This ranking is remarkably robust: it is unchanged under alternative policy regimes (a higher inflation target, a negative lower bound, average inflation targeting, a dual mandate), alternative assumptions about the underlying trend, alternative learning speeds, and alternative household forecasting models.

Second, a sufficiently large policy response to inflation combined with a small response to the OG estimate—as in the original Taylor rule coefficients—robustly delivers good welfare performance, regardless of the detrending method. We temper this positive observation by noting that a large response to the OG estimate can be bad for welfare even for monetary policies that satisfy the celebrated Taylor principle.

Third, the differences among methods are state dependent. They are particularly stark during severe recessions or crises when the ZLB binds; in normal times, when the ZLB does not bind, stability regions are broadly comparable across methods. Particular caution in the choice of method is therefore warranted in times of crisis. Relatedly, policy regimes that relax the ZLB constraint—a higher inflation target, a negative lower bound, average inflation targeting—enlarge stability regions and raise welfare for every method.

Fourth, while the optimal response to inflation is similar in magnitude across methods,

[2] See similar discussions in Christiano and Gust (1999) and Bullard and Mitra (2007), albeit in different contexts.

[3] By 'intrinsically' we mean the properties of the estimates a method produces in the typical setting of econometricians, i.e., without feedback from the choice of estimation method to the data generating process.

the optimal response to the OG estimate is not: methods which intrinsically produce large and volatile OG estimates, such as polynomial detrending, call for a small reaction coefficient, whereas methods producing smaller and less volatile estimates, such as the bandpass and BN-KMW filters, call for a substantially bigger one.[4]

Fifth, we connect stability to the "reliability" criterion of Orphanides and van Norden (2002), who evaluate estimation methods by the size of their real-time revisions while treating the data generating process as exogenous. In our model the data generating process is endogenous to the method choice, yet a number of their findings survive this dynamic interaction. More importantly, stability and reliability turn out to be closely associated: methods with large stability regions also produce small revisions.

The paper bridges the econometric literature on OG estimation and the macroeconomic literature on monetary policy design, and its questions have not been addressed in either. In the policy-design literature, the seminal contribution of Bullard and Mitra (2002) examines Taylor-type rules that lead to "learnable" equilibria, and a related strand advocates rules that avoid indeterminacy under rational expectations (RE), e.g., Bernanke and Woodford (1997). Throughout, the central bank is assumed to observe the OG entering its rule; NK models in the full-information RE tradition of Woodford (2003) and the imperfect-knowledge tradition of Orphanides and Williams (2007, 2008) and Eusepi and Preston (2018) likewise treat the OG measure as given. We instead make the measurement of the OG—the choice of detrending method—the object of analysis, and policies desirable under our stability criterion often differ from those selected by learnability or determinacy; even when they coincide, the reasons for their desirability differ substantially. On the econometric side, a large literature compares the properties of OG estimation methods under an exogenous data generating process, e.g., Orphanides and van Norden (2002) and Canova (2025); our contribution is to endogenize the data generating process through the policy feedback loop. The eleven methods we consider range from the most heavily used to recently developed ones, include purely statistical as well as theoretically founded methods, and reflect the small-model practice favored in the estimation literature, e.g., Kamber et al. (2018), Hamilton (2018) and Coibion et al. (2018).

The remainder of the paper is organized as follows. Sections 2–4 present the model, learning by households and the policymaker, and the parameterization and simulation of the model. Sections 5 and 6 contain the baseline results on stability, optimal policy and the reliability of

[4] An interest rate rule simply responding to inflation and the growth rate of output, like the one in Orphanides and Williams (2007) (i.e., the FOD method, which does not require estimating the OG), has good stability and welfare properties in our model.

detrending methods. Section 7 summarizes extensions and robustness checks, with full details relegated to the appendix, and Section 8 concludes.

2 Model

We employ a structural New Keynesian framework that approximates the US economy to analyze learning about the OG and decisions by the monetary policymaker. The framework is closely based on Woodford (2003) and Clarida et al. (1999). Similar model specifications have been widely utilized in the literature, e.g. Swanson and Williams (2014), Smets (2003), Orphanides (2002) and Orphanides and Williams (2007, 2008), in the context of monetary policy design.[5]

2.1 The structural model

The "IS equation" (2.1) is motivated by the consumption Euler equation and describes the behavior of the OG:

$$x_t = \alpha x^e_{t+1} + \phi x_{t-1} + \eta(i^e_t - \pi^e_{t+1}) + e_{x,t}, \tag{2.1}$$

$$e_{x,t} = \rho_x e_{x,t-1} + \upsilon_{x,t}, \upsilon_{x,t} \sim IIN(0, \sigma^2_x), \tag{2.2}$$

where IIN means independent and identically normally distributed. Output gap x_t is calculated as the difference between (log) output and (log) natural rate of output (or potential output) and is given by $x_t = y_t - y^p_t$, where y_t is (log) output at period t, and y^p_t stands for the true (log) potential output. The trend labor productivity process a^p_t is assumed to be a random walk with drift process (2.3). So the evolution of y^p_t is determined by the stochastic trend productivity process

$$\Delta a^p_t = g + e_{a,t}, \tag{2.3}$$

$$\Delta y^p_t = \Delta a^p_t, \tag{2.4}$$

where g is the long-run trend growth rate and $e_{a,t} \sim IIN(0, \sigma^2_a)$ is a technology shock. i_t denotes the deviation of short-term nominal interest rate from its long-run equilibrium level and π_t denotes the inflation rate in period t (first difference of the log of the price level). The demand shock $e_{x,t}$ is modeled as an AR(1) process to capture potential serial correlation. $(.)^e$ is the subjective expectation operator that may or may not coincide with rational expectations (RE).

[5]The literature analyzes monetary policy rules in a wide variety of models, from small to large scale and ad hoc to optimizing (see Taylor, 1999, especially the Introduction, and Levin et al., 2003). Since we address a novel question, we start with a simple model inspired by Orphanides and Williams (2007, 2008); our results are robust to alternative models, including the fully micro-founded infinite horizon learning formulation of Preston (2005, 2006) (Appendix P).

The IS relation (2.1) describes the relationship between OG and its determinants, OG forecast x^e_{t+1}, lagged OG and ex ante real interest rate forecasts, $(i^e_t - \pi^e_{t+1})$.[6] α denotes the importance of expectation in determining OG and ϕ reflects OG inertia.

The random walk with drift specification for potential output in equations (2.3)–(2.4) is standard in the literature (e.g., Orphanides and Williams, 2007, 2008; Laubach and Williams, 2003; Rotemberg and Woodford, 1996) and empirically well-supported for US real GDP (Nelson and Plosser, 1982). It is consistent with the stochastic trend representation underlying several of the methods we consider (e.g., the Beveridge-Nelson decomposition, the Hamilton filter and first-order differencing), while deterministic detrending methods and the HP filter do not assume a stochastic trend and are thus potentially misspecified. Evaluating how methods perform when possibly misspecified relative to the true data generating process is, however, precisely part of our goal, and the relative ranking of methods is robust across alternative specifications of the trend, including a deterministic trend and a trend with shocks to its growth rate (Section 7), which addresses documented shifts in US trend growth (e.g., Perron and Wada, 2009; Luo and Startz, 2014).

The "Phillips curve" is motivated by the new Keynesian Phillips curve (NKPC),

$$\pi_t = \beta \pi^e_{t+1} + \kappa x_t + e_{\pi,t}, \tag{2.5}$$

$$e_{\pi,t} = \rho_\pi e_{\pi,t-1} + v_{\pi,t}, v_{\pi,t} \sim IIN(0, \sigma^2_\pi). \tag{2.6}$$

The cost-push shock $e_{\pi,t}$ is modeled as an AR(1) process. The Phillips curve can be derived from the monopolistically competitive firms' profit maximization problem under Calvo (1983) staggered price setting. It relates current inflation (measured in the data as the annualized percentage change in the real time GDP price index, or GNP before 1991:Q4) to inflation forecast and the contemporary OG. β is interpreted as the discount factor of firms.

To match the lag in information structure (Woodford, 2003; Giannoni and Woodford, 2004; Orphanides and Williams, 2007) in period t, households are assumed to form their forecasts on the information available at date $t-1$ (or equivalently the relevant forecasts are formed in the previous quarter). So private agents cannot observe current output gap, inflation and interest rate before making decisions in period t. Based on their forecasts and values of pre-determined variables, current inflation, output gap and interest rate are determined simultaneously through equations (2.1), (2.5), and (2.7) below.

[6] In common with the vast majority of the literature on monetary policy, we assume trend inflation to be zero. It would, of course, be of interest to consider the case of positive trend inflation, as say in Ascari and Sbordone (2014), but this would make the analysis substantially more complicated and we leave this for future work.

2.2 Monetary policy and the loss function

The NK model is completed by including a monetary policymaker who is responsible for interest rate decisions in each period, as in e.g., Bullard and Mitra (2007) and Williams (2006). Following most of the literature, the central bank commits to a simple interest rate rule (2.7) that incorporates a ZLB on nominal interest rates:

$$i_t = \max\{\theta_r i_{t-1} + \theta_\pi \pi_{t-1} + \theta_x \widetilde{x}_{t-1} + e_{i,t}, -(\pi^* + r^*)\}, \tag{2.7}$$

where $e_{i,t} \sim IIN(0, \sigma_r^2)$ is the monetary policy shock. π^* is the inflation target set by the policymaker and r^* is the long-run natural rate of interest.[7] $\widetilde{x}_{t-1}$ is the policymaker's real-time estimate of the OG obtained by a detrending method using information available up to period $t-1$; $\widetilde{x}_{t-1} = y_{t-1} - \tilde{y}^p_{t-1}$, where $\tilde{y}^p_{t-1}$ is the estimated potential output. The "max" function in rule (2.7) reflects the presence of the ZLB on nominal interest rates. While several central banks have experimented with negative policy rates in recent years, we adopt a strict zero bound in the baseline model for simplicity and consider a negative lower bound as an extension in Section 7. The central bank adjusts interest rates in response to lagged inflation and estimated OG from the previous period, reflecting a possible lag in information release as evidenced in reality.[8]

Loss function. The performance of monetary policies is evaluated using a standard widely used loss function, e.g., Orphanides and Williams (2007). The loss is calculated as the weighted sum of unconditional volatilities of inflation and OG:

$$L = Var(\pi) + \lambda_x Var(x), \tag{2.8}$$

where $Var(a)$ is the unconditional variance of variable a and λ_x is the relative weight the policymaker places on OG volatility. This criterion is closely connected to central banks' objectives in practice: the Fed's dual mandate of "maximum employment and price stability" can be operationalized as a loss function penalizing volatility in inflation and economic slack (Yellen, 2012). Reflecting the predominant focus on inflation stabilization among inflation-targeting central banks (Woodford, 2003; Svensson, 2010; Blanchard and Galí, 2007) and following Giannoni and Woodford (2004), we choose a small relative weight on OG variability, $\lambda_x = 0.003$, in the baseline. Section 7 considers alternative regimes: a higher inflation target, a dual mandate (Yellen, 2012; Debortoli et al., 2019), average inflation targeting (Brainard, 2019; Svensson, 2020), and a slightly

[7] Note that the natural rate of interest is time-varying in the model, but we assume that only its long-run mean value is incorporated into the policy rule.

[8] An alternative assumption is for interest rates to respond to year-on-year inflation instead of quarterly inflation. We conjecture such an alternative would not change the main messages of the paper.

negative lower bound.

Information set and belief formation. We make several remarks regarding the information set and belief formation by households and the policy maker. First, the policy maker is assumed to have imperfect knowledge about the economy and face uncertainty regarding potential output and the output gap (OG). In particular, we assume that the policymaker does not have RE. She knows the productivity process (equation (2.3)) but not the potential output process (equation (2.4)) due to imperfect knowledge of, for instance, households' preferences or beliefs. She is assumed to use one of the estimation methods explained in Section 3.2 to estimate the OG each period. Her subjective estimates of potential output can deviate from the objective potential output (equation (2.4)), indicating a deviation from RE.

Second, and importantly, there is no information asymmetry between households and the policymaker: neither has RE or knowledge of the potential output process (equation (2.4)). Households do not need this knowledge to make decisions; they only forecast the interest rate, inflation, and the growth rate of output, as explained in Section 3.1. We acknowledge an asymmetry in sophistication—households learn about simple univariate processes while the policymaker uses more sophisticated econometric methods to estimate the OG—which reflects the reality that central banks employ teams of economists with access to advanced models and rich datasets, whereas individual households typically rely on simple heuristics. Our results are robust to alternative specifications of households' forecasting models, including correctly specified models that nest the RE solution (Appendix O).

3 Expectations and learning

This section introduces learning and expectation formation by households and the policymaker. We adopt adaptive learning because the RE assumption of complete knowledge of the economy is particularly strong in our context: if the central bank itself cannot observe the OG and must estimate it in real time, it is natural that private agents also face imperfect knowledge (Eusepi and Preston, 2018). Adaptive learning also generates realistic expectation dynamics consistent with survey evidence on forecasts, and it sustains the feedback loop between OG estimation, monetary policy and aggregate outcomes that is central to the paper—a loop absent under RE, where the economy's law of motion is known.

3.1 Learning by households

Households are assumed to form expectations by adaptive learning, as in Marcet and Sargent (1989) and Eusepi and Preston (2018). Each period, households need to forecast output, inflation and interest rate to make economic decisions. In the baseline scenario, they are assumed to use simple models for forecasting:

$$\pi_t = \gamma^{\pi} + \epsilon_t^{\pi}, \tag{3.1}$$

$$i_t = \gamma^{i} + \epsilon_t^{i}, \tag{3.2}$$

$$\Delta y_t = \gamma^{\Delta y} + \epsilon_t^{\Delta y}, \tag{3.3}$$

where ϵ_t^{π} , ϵ_t^{i} and $\epsilon_t^{\Delta y}$ are perceived to be $i.i.d.$ and mutually independent. Thus, households believe that inflation and interest rates are a constant plus some i.i.d. noise (often called steady state learning in the literature) while output growth is a constant plus i.i.d. noise.[9]

Households recursively update estimates of γ^{π}, γ^{i} and $\gamma^{\Delta y}$ via a constant gain learning algorithm:

$$\gamma_t^{w} = \gamma_{t-1}^{w} + k^{w}(w_t - \gamma_{t-1}^{w}), \tag{3.4}$$

where $k^{w} > 0$ is the gain parameter and $w = \pi$, i or Δy. This parsimonious specification has several advantages in forecasting. The forecasting literature finds under-parametrized models produce more accurate short-run forecasts than complicated ones with small samples, so that households may optimally adopt a simple model in forecasting, e.g., Fuster et al. (2010). Moreover, beliefs related to steady state learning fit reality closely and are generally supported by many studies using experiments, e.g., Adam (2007) and Armona et al. (2019).[10] Like the policymaker, at period *t*, households also use information up to period *t-1* to form expectations, as in Eusepi and Preston (2011) and Adam et al. (2012). Later on, we show that our main results are robust to alternative assumptions on households' expectation formation including when households' forecasting model nests the RE solution (i.e., the model is correctly specified in learning literature parlance and analyzed in Appendix O).

We note that in the current framework, households do not learn about the policy rule or the policymaker's estimation method. A large body of survey evidence suggests that households have limited awareness of monetary policy rules and central bank actions (e.g., Coibion et al., 2022),

[9] We do not assume that agents have knowledge of the balanced growth path, consistent with survey evidence in Kuang et al. (2022, 2025).

[10] In adaptive learning parlance, these beliefs are under-parameterized relative to the RE solution; see Di Pace et al. (2021) for how under-parameterized beliefs can improve the fit of the search and matching model.

and of the monetary transmission mechanism: central bank communication about interest rates alone has little effect on households' house price expectations unless the mortgage rate channel is also explained (Binder et al., 2026). In addition, the constant-gain specification implies that agents perpetually weigh recent observations more heavily, which can be interpreted as reflecting alertness to possible structural changes; this approach is standard in the adaptive learning literature (see Eusepi and Preston, 2018, for a discussion).

3.2 Learning by the policymaker

A fundamental issue faced by the policymaker is that the OG has to be estimated since it cannot be observed. The policymaker makes use of historical data and adopts a detrending method to estimate the OG.[11] The resulting real-time OG estimate is used as an input for interest rate decisions according to equation (2.7). In the next period, when a new data observation becomes available, the policymaker revises the estimate of the OG.

We consider eleven widely used methods advocated in the literature and adopted in practice, including *linear and quadratic trend, Discounted Least Squares (DLS) detrending, first-order differencing (FOD), the (one-sided) Hodrick-Prescott filter* (Stock and Watson, 1999), *Hamilton filter* (Hamilton, 2018), *Christiano-Fitzgerald band-pass filter* (Christiano and Fitzgerald, 2003), *Beveridge-Nelson decomposition* (Beveridge and Nelson, 1981), *Beveridge-Nelson decomposition with a low signal-to-noise ratio* (Kamber et al., 2018), *multivariate Beveridge-Nelson decomposition* and *Blanchard-Quah filter* (Blanchard and Quah, 1989; Coibion et al., 2018).

These detrending methods generally decompose (log) output y_t into a secular component (trend) and a cyclical component (cycle):

$$y_t = y_t^{tr} + \widetilde{x}_t, \tag{3.5}$$

where y_t^{tr} is the trend and $\widetilde{x}_t$ is the cycle. In the baseline monetary policy rule (2.7), the policymaker considers the trend component as the real time estimate of potential output $\widetilde{y}_{t-1}^{p}$ and the cyclical component as the real time estimate of OG $\widetilde{x}_{t-1}$. The use of the lagged estimate of the OG reflects the lag in information release.

We acknowledge that in practice, modern central banks do not mechanically apply a single statistical filter: no central bank relies on any single method, and large teams of economists combine multiple models, judgment and various indicators beyond GDP (Edge and Rudd, 2016). Our assumption that the policymaker uses a single estimation method is a deliberate simplification

[11] Relatedly, Kuang and Mitra (2016) study learning about long-run growth rates in a real business cycle model, and Kuang and Mitra (2025) study the role of mis-measured OG in EU fiscal policy after the 2007-08 crisis.

that transparently isolates how the properties of each method interact with monetary policy and aggregate outcomes. This is similar in spirit to the literature that evaluated commitment to simple Taylor-type rules across wide varieties of NK models, where it was likewise recognized that no central bank literally commits to a simple rule (see, for example, the volume edited by Taylor, 1999, and Levin et al., 2003).[12]

We now provide a brief explanation of the OG estimation methods.[13] Appendix A provides a detailed explanation of these methods. The OG estimates from linear and quadratic detrending methods are the residuals after removing a linear and quadratic time trend from the output data, respectively. For example, the U.S. Congressional Budget Office (CBO) continues to use linear detrending to estimate potential output and the output gap. With the DLS detrending method, output data is first discounted at a constant rate, $0 \leq d \leq 1$. The OG estimates are the residuals obtained by removing a linear time trend from the discounted output series $\{d^{T-t} \cdot y_t\}_{t=1}^{T}$. The discount factor is set to $d = 0.99$.

The FOD method assumes the trend equals the level of output from the previous period. Thus the OG is the growth rate of output. The one-sided HP filter with a smoothing parameter of 1600 is used, following Stock and Watson (1999) and Orphanides and van Norden (2002). The OG estimates produced by the filter of Hamilton (2018) are the deviations of output from the fitted value from a regression of output on four lagged values of output back-shifted by two years and a constant. The bandpass filter of Christiano and Fitzgerald (2003) isolates the cyclical component of output whose period of oscillation falls within a chosen range (w^l, w^u) with $w^l = 1.5$ years and $w^u = 8$ years.

The cyclical component of output produced by the filter of Beveridge and Nelson (1981) is the difference between current output and the value it is expected to have in the indefinite future, abstracting from the unconditional mean of the growth rate of output. Following Morley et al. (2003), the BN filter with an AR(2) model for output growth is used. The BN-KMW filter is a modified BN filter and implemented following Kamber et al. (2018), estimating an AR(12) model and imposing a low signal-to-noise ratio 0.05. The multivariate BN filter makes use of output growth data and unemployment rate to estimate trend output rate; eight lags are used in the VAR.[14] Following Coibion et al. (2018), structural VAR models of output growth and

[12] Extensions incorporating model averaging, judgment, or method-switching based on performance would be natural directions for future research.

[13] Note that in real-time analysis, the OG estimate of interest is at the end of the observation sample. This precludes the policymaker's adoption of some two-sided (or symmetric) filters that use an equal number of observations before and after period t, to isolate the trend or cycle at period t, see van Norden (2002).

[14] A 2-variable VAR with output growth and inflation as in Bullard and Keating (1995) is not suitable for estimating

unemployment rates with the Blanchard and Quah (1989) identification scheme and eight lags are used to produce estimates of potential output (the BQ filter). [15] To apply the multivariate BN filter and the BQ filter, we require unemployment data as an input. In these cases, the three-equation NK model is supplemented with Okun's law, using a coefficient of 2.

4 Parameterization and simulation

4.1 Calibration and estimation

We closely follow the method in Orphanides and Williams (2007) to obtain parameter values. Some parameters are calibrated, while others are estimated using survey forecasts and realized data. The trend productivity growth g is calibrated to be the average trend output growth estimated using the methodology of Holston et al. (2017) and the sample period 1971:Q1 to 2019:Q4. The long-run natural rate of interest r^* is set to the historical average natural rate of interest. In particular, g is set to 2.88% per year and annualized r^* to 4% per year. The baseline (annual) inflation target π^* is set to 2%. Alternative and higher inflation targets are also considered later. Both the standard deviation of technology shock and monetary policy shock are set to 0.5%, consistent with the estimates in the literature, e.g., Ireland (2004) and Davig and Doh (2014).

To estimate the IS curve (2.1), a small weight is placed on lagged OG. In particular, we set $\alpha = 0.95$ and $\phi = 0.02$. In RE models, a large coefficient on lagged OG is typically needed to capture the dynamics of OG in the data. However, a small coefficient on lagged OG data suffices in the learning model as learning generates the degree of persistence in the data, see, e.g., Milani (2007).[16] OG forecasts are generated via adaptive learning in the model, which helps to reproduce realistic OG dynamics without a large coefficient on lagged OG. Survey forecast data is used to estimate the coefficient on how OGs respond to interest rate (η) and the parameters of the shock process (ρ_x, σ_x).

In the Phillips curve (2.5), the discount factor β is set to 0.99. The remaining parameters

the OG: both of its shocks have a permanent effect on output, leaving no distinction between trend and actual output.

[15] Methods using inflation to estimate the OG (e.g., Holston et al., 2017) are computationally very demanding when the policymaker learns recursively; we leave them for future work.

[16] Milani (2007) estimates the degree of habit formation to be in the range of [0.006, 0.289], see his Table 2, p. 2073. Our choice of the degree of habit formation (0.02) is consistent with Milani's estimate. In addition, here the sum of α and ϕ is less than 1, consistent with e.g. Gabaix (2020) and Budianto et al. (2023). Gabaix motivates this feature based on cognitive discounting or cognitive inattention on the part of households/firms. Our results are, nevertheless, robust to alternative parameterizations, e.g., $\alpha = 0.5$ and $\phi = 0.5$; see Appendix O.

$(\kappa, \rho_\pi, \sigma_\pi)$ are estimated using survey forecast data. The adaptive learning model does not rely on a lagged inflation term as the survey forecasts of inflation rates are highly autocorrelated and learning generates this persistence. Due to the unavailability of OG forecast data, we first estimate an unemployment gap version of the model and then apply Okun's (1962) Law to bridge the OG and unemployment rate to recover estimates for the OG model in specification (2.1) and (2.5).[17] The realized unemployment rate and inflation are taken from the Real-Time Data Set for Macroeconomists. The median 1-quarter-ahead forecasts of unemployment rate and three-month Treasury bill rates from the Survey of Professional Forecasters (SPF) are chosen as proxies for expected unemployment and short-term nominal interest rate, respectively. Both datasets are managed by the Federal Reserve Bank of Philadelphia. We construct inflation forecasts by computing the annualized log difference of GDP or GNP price index forecast from SPF. For estimating the unemployment equation, we use the most recent estimates of Natural Rate of Unemployment (NAIRU) retrieved from Congressional Budget Office (2020) as the actual historical natural rates of unemployment. When estimating, we subtract the most recent estimates of natural rate of interest, obtained from Holston et al. (2017), from forecasts of Treasury bill rate in the sample.

Using an Okun's coefficient of 2, Appendix C estimates the IS curve as follows.

$$x_t = 0.95 x^e_{t+1} + 0.02 x_{t-1} - \underset{(0.104)}{0.296}(i^e_t - \pi^e_{t+1}) + e_{x,t}, \tag{4.1}$$

$$e_{x,t} = \underset{(0.068)}{0.390} e_{x,t-1} + \upsilon_{x,t}, \qquad \sigma_x = 0.58, \tag{4.2}$$

$$\pi_t = 0.99\pi^e_{t+1} + \underset{(0.011)}{0.027} x_t + e_{\pi,t}, \tag{4.3}$$

$$e_{\pi,t} = \underset{(0.069)}{0.360} e_{\pi,t-1} + \upsilon_{\pi,t}. \qquad \sigma_\pi = 0.41. \tag{4.4}$$

The constant gain parameters, k^π, k^i and $k^{\Delta y}$ in households' belief updating are calibrated by minimizing the distance between the forecasts generated by adaptive learning and the corresponding survey forecasts. This minimization exercise yields the gain parameters, $k^\pi = 0.008$, $k^i = 0.004$ and $k^{\Delta y} = 0.003$, see Appendix D. These values generally fall within the appropriate range of parameterization, 0.002 - 0.05, employed in the literature; see e.g., Eusepi and Preston (2011) and Kuang and Mitra (2016).[18]

[17] Following the literature (e.g., Okun, 1962; Knotek, 2007; Orphanides and Williams, 2007; Yellen, 2012), Okun's law, $u_t = -cx_t + u^*$, is used with a coefficient of 2, where u_t is the unemployment rate. u^* is the natural rate of unemployment (NAIRU) and set to be equal to the average of the estimates made by the Congressional Budget Office during 1971Q1-2019Q4 (5.5%).

[18] For instance, in Eusepi and Preston (2011), households need to forecast interest rates and they calibrate the gain parameter to be 0.002.

4.2 Model solution and simulation

Model solution. The temporary equilibrium in the model with a ZLB under learning can be solved easily. Due to the lagged information structure and time - $t-1$ dating explained in Section 2.1, expectations of interest rates, inflation, and the OG are predetermined. Combined with the historical realizations of variables, equation (2.1) directly provides the solution for the OG and hence the current output level. Given the current output level and the historical realizations of variables, equation (2.7) determines the nominal interest rate. Finally, using the OG and the predetermined inflation expectations, equation (2.5) determines the inflation rate.

Model simulation. We evaluate the performance of monetary policy based on the loss function (2.8). We assume that the central bank commits to the interest rate rule (2.7) and chooses the policy coefficients, θ_r, θ_π and θ_x. In all cases, stochastic simulations are implemented to compute approximated unconditional moments under various combinations of policy coefficients. The combination of policy coefficients that minimizes the loss function (8) is referred to as the "optimal monetary policy".

Computational details. For each combination of policy parameters on the grid, we perform stochastic simulations and compute the welfare loss. The grid for policy coefficients uses increments of 0.1 for both θ_π and θ_x. For each grid point, the model is solved as a temporary equilibrium system (as described above), where the predetermined expectations and lagged variables yield a unique solution for current-period variables in each period. No iterative solver or convergence criterion is required for the within-period solution since the system is recursive given the lagged information structure. The optimal policy is identified as the grid point achieving the minimum welfare loss. All simulations are performed using MATLAB and the computations were carried out using the University of Birmingham's BlueBEAR high performance computing service.

Every period, the policymaker uses the most recent T periods of data to form the estimate of OG. We set T =120 periods (30 years).[19] We simulate our model for 1,000+40,000 periods. The first 1,000 periods are discarded in order to mitigate the impacts of initial conditions. The final 40,000 periods are utilized to calculate associated loss and other relevant statistics, such as revisions to OG estimates and their autocorrelation coefficients. In each simulation, all model variables are initialized to their corresponding long-run growth or steady-state levels. Households' initial beliefs are likewise fixed to the RE values: initial beliefs about the parameters

[19] A rolling-window of 120 quarters of data is also used in Coibion et al. (2018) to estimate the OG.

in the forecasting model of interest rate and inflation are set to be 0 (being deviations from the steady state), whereas the initial belief about long-run output growth is g. In each period, a vector of innovations, $[v_{x,t}, v_{\pi,t}, e_{a,t}, e_{i,t}]$, is drawn from *IID* Gaussian distributions with estimated (or calibrated) variances as reported. The policymaker updates the OG estimate, $\widetilde{x}_{t-1}$, using a rolling window of the latest T output observations $\{y_{t-j}\}_{j=1}^{T}$ collectable at period $t-1$, and sets the interest rate accordingly. Then the structural model is simulated, households' beliefs are updated, and households' new forecasts of inflation, interest rate, and output are obtained.[20] In the presence of ZLB on nominal interest rates, the relevant literature (Williams, 2006) also imposes ZLB on the agent's forecasts of interest rate. In our baseline model, agents' beliefs (3.2) and the constant-gain learning rule (3.4) provide a natural ZLB on their forecasts of the interest rate: as actual interest rates are bounded below by zero, their forecasts are also bounded below by zero. In addition, when implementing multivariate BN decomposition and BQ sVAR, the model uses Okun's law to generate unemployment in each period as an input for the filters.

5 Baseline results: stability

This section analyzes the welfare properties of the combined choice of a detrending method (out of eleven methods) and the consequent interest rate setting by the policymaker. Section 5.1 summarizes the stability properties of the methods while Section 5.2 discusses the stability properties of the (famous) Taylor rule and Section 5.3 explores the underlying economic mechanism driving the results.

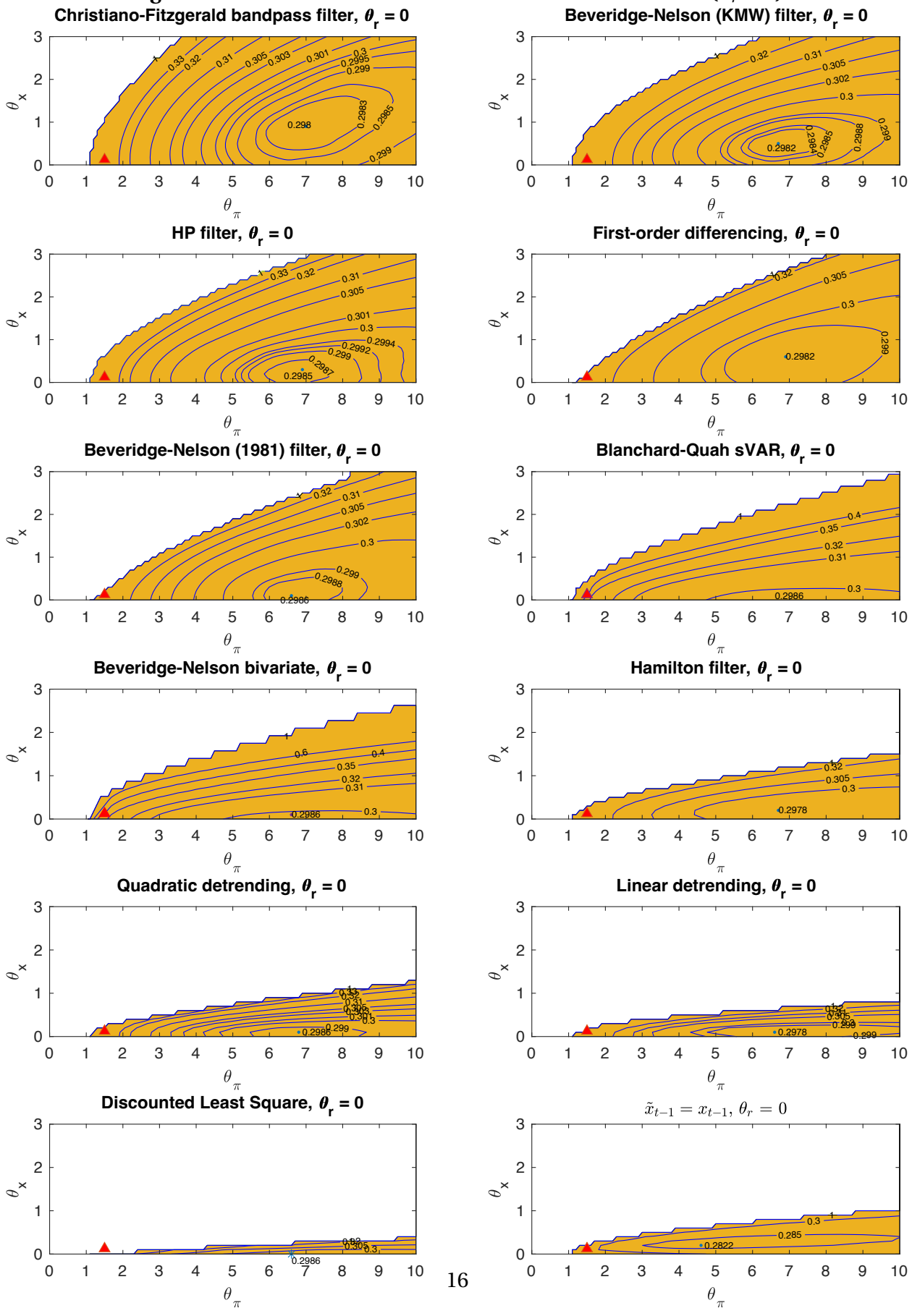
5.1 The iso-loss contours and stability region

Figure 1 displays iso-loss contours associated with eleven detrending methods, when there is zero degree of inertia ($\theta_r = 0$) in interest rates.[21] Appendix E illustrates iso-loss contours for three alternative degrees of inertia: $\theta_r = 0.2$, 0.6, and 0.8. For the sake of comparison, we also consider the scenario when the policymaker has knowledge of the true OG while agents continue to revise their beliefs over time in the economy. This scenario is shown by the panel on the bottom right. In each panel, the horizontal axis shows the (policy) reaction coefficient of inflation (θ_π) and the

[20] Households' two-quarter (or N-quarter) ahead OG forecast is derived in Appendix B. The adaptive learning literature often augments the belief updating rule with a "projection facility", to avoid explosive behavior of beliefs and economic variables, as in Orphanides and Williams (2007, 2008) and Marcet and Sargent (1989). The results reported in our paper do not make use of a projection facility.

[21] In this and all the figures in the paper, we display the welfare loss for a wide range of policy parameters in the tradition of the vast literature on monetary policy rules, see e.g. Rotemberg and Woodford (1999) and Woodford (2003).

Figure 1: ISO-LOSS CONTOURS WITH DETRENDING METHODS ($\theta_r = 0$)



vertical axis the reaction coefficient of OG (θ_x). The dot in each panel depicts the minimum loss achieved in each scenario. From a normative perspective, the policymaker would like to avoid detrending methods (alternately OG estimation methods) which lead to volatile or unstable paths of the endogenous variables since this would in turn lead to large welfare losses due to the policymaker's loss function. This normative consideration motivates our stability criterion which is based on welfare loss above a threshold value. The welfare loss is determined by the volatilities in inflation and OG and provides a summary measure for this criterion.[22] Some detrending methods lead to stability over a wide choice of policy parameters whereas others do so only over a narrow choice of parameters. As explained in the Introduction, we view detrending methods with a larger stability region more desirable. This in turn provides an obvious criterion to rank different methods based on the size of their stability region.

We set the stability threshold at a welfare loss of one. Losses escalate sharply at the edge of the stable region—the iso-loss contour for a loss of 100 nearly coincides with that labeled "1"—so the ranking of methods is insensitive to the particular threshold. In each panel, the shaded region is stable (welfare loss below one) and the unshaded region, outside the contour labeled "1", is unstable.[23]

To avoid instability, it is necessary to have a sufficiently strong response to inflation; this is true for all detrending methods and all degrees of interest rate inertia. For instance, for the linear detrending method in Figure 1, the left-hand-side iso-loss line (labelled as "1") in the figure coincides with the horizontal axis (θ_π value) at around $\theta_\pi = 1$. The model exhibits explosive behavior if $\theta_\pi < 1$ for all methods when there is no interest rate inertia. This resembles the idea underlying the Taylor principle that the nominal interest rate should respond to inflation "more than one-for-one" to stabilize the economy. However, note that too large a response to the OG estimate will generally lead to instability, even for policies satisfying the Taylor principle (a phenomenon we explain in the next section).

The OG estimation methods are associated with different sizes of the stable region. The size of the stable region is calculated as the ratio of the area of the parameter space associated with a loss less than 1 relative to the area of the full parameter space, i.e., $\theta_x \in [0,3]$ and $\theta_\pi \in [0,10]$. Panel A of Table 1 reports the size of the stable region for the detrending methods with different degrees of interest rate inertia. We summarize the results below.

[22] Alternately, one may provide thresholds for the volatilities in the individual variables or even directly on the behavior of variables like inflation, OG (or their expectations). The welfare loss criterion is appealing since this is, after all, the primary concern of the central bank in our setup.

[23] The contour plots have some discontinuities at the border of stability and instability region. This is because of the coarse grid of θ_π and θ_x.

Table 1: THE SIZE OF STABILITY REGION FOR DIFFERENT DETRENDING METHODS (%)

Panel A: with ZLB	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	79.8	80.5	80.3	78.5
Beveridge-Nelson (KMW, 2016)	70.6	70.6	69.3	67.6
HP filter	66.4	66.1	63.0	59.4
First-order differencing	58.8	61.2	64.6	65.4
Beveridge-Nelson (1981)	56.4	58.6	61.9	62.9
Blanchard-Quah	53.7	45.5	42.7	34.6
Beveridge-Nelson (bivariate)	50.0	40.1	38.0	29.7
Hamilton	28.7	28.2	25.9	23.6
Quadratic trend	23.2	22.7	20.5	15.4
$\widetilde{x}_{t-1} = x_{t-1}$	20.0	19.1	15.0	8.3
Linear trend	16.3	15.3	12.2	9.5
Discounted Least Squares	1.7	1.8	1.7	1.6
Panel B: without ZLB				
Christiano-Fitzgerald	90.0	92.0	96.0	98.0
Beveridge-Nelson (KMW, 2016)	90.0	92.0	96.0	98.0
HP filter	90.0	92.0	96.0	98.0
First-order differencing	88.6	90.7	95.0	97.0
Beveridge-Nelson (1981)	88.6	90.7	95.0	97.0
Blanchard-Quah	90.0	92.0	96.0	92.9
Beveridge-Nelson (Bivariate)	90.0	92.0	96.0	72.8
Hamilton	88.2	91.1	94.9	83.3
Quadratic trend	90.0	92.0	48.0	21.8
Linear trend	90.0	92.0	60.6	39.0
Discounted Least Squares	46.7	47.3	43.9	36.4

Notes: The table reports the size of the stability region for each detrending method under different degrees of interest rate inertia (θ_r). The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. Panel A imposes the zero lower bound (ZLB) on nominal interest rates; Panel B removes the ZLB constraint. The interest rate rule is $i_t = \max\{\theta_r i_{t-1} + \theta_\pi \pi_{t-1} + \theta_x \widetilde{x}_{t-1} + e_{i,t}, -(\pi^* + r^*)\}$. Methods are ranked by the size of their stability region under $\theta_r = 0$ in Panel A.

The CF band-pass filter and BN filter with a low signal-to-noise ratio (BN-KMW) have the largest stability regions (which is also robust to different degrees of interest rate inertia). The HP filter (commonly used by policy making institutions) and the BQ filter perform well as they are associated with large stability regions. The two multivariate detrending methods, bivariate BN filter and the BQ filter, are associated with large stability regions but underperform relative to some univariate detrending methods, such as the BN-KMW filter and the bandpass filter.

On the other hand, the stable region associated with DLS detrending is the smallest (between 1.6% – 1.8% across different degrees of interest rate inertia). For both linear and quadratic detrending, the stable regions are small no matter what the response to the lagged interest rate is.

Consider now the economy where the policymaker has knowledge of the true OG while agents continue to learn over time. Interestingly, in this case, the stable region is small and in fact smaller than almost all detrending methods (except for DLS and linear detrending methods), see the bottom right panel of Figure 1.[24] Thus, accurately tracking the OG may not be the best thing to do when there is a dynamic feedback between output gap estimation, interest rate decisions, and aggregate outcomes.[25]

Because our evaluation lets the chosen method itself change the data generating process, the desirability of detrending methods can differ from that in earlier studies, as the following comparison illustrates.

Canova (2025) favors polynomial detrending and FOD filter in extracting gaps and transitory components. We, however, find that linear, quadratic and DLS detrending methods are associated with small stable regions and perform poorly in the baseline New Keynesian model. So if any of these deterministic detrending methods is adopted by the policymaker, the choice of policy coefficients is delicate since they often lead to instability. This will be a reason for concern owing to the reasons explained in the Introduction. The FOD filter favored by Canova (2025), on the other hand, performs better than deterministic detrending methods and is associated with a large stable region (though it underperforms relative to the BN-KMW filter and CF bandpass filter).[26]

[24] Here the stability region is small because the true OG estimate is large and volatile which can easily lead to large deflation and instability in the presence of the ZLB, see the last paragraph of Section 6.1 for further explanation.

[25] In our analysis, the Taylor rule responds to lagged values of inflation and the OG estimate. Many central banks in practice use forecast-based rules where interest rates respond to projections of future inflation and economic activity. Under such forward-looking rules, the interaction between OG estimation methods and expectations could be amplified, as forecast errors would feed back into both the policy rate and agents' expectations simultaneously. Analyzing forecast-based Taylor rules in our framework is an important extension for future research.

[26] We remark that, unlike in Orphanides and Williams (2007), the differenced rule ($\theta_r = 1$) is associated with a small

Full-information RE benchmark. Readers may wish to have a benchmark to compare the outcomes of different OG estimation methods. One benchmark is the full-information RE case where all agents including the policymaker have complete information and form RE. Figure A.12 in Appendix H displays the iso-loss contours and the optimal policy under full-information RE. The optimal policy under full-information RE achieves a lower welfare loss than under learning because learning propagates and amplifies economic shocks and reduces the policymaker's ability to stabilize the economy. Additionally, virtually all choices of policy parameters do not lead to very large welfare losses under full-information RE.[27]

Alternative loss function. In the above analysis, the policymaker's loss function depends on the volatility of the true output gap. In the instability region, the choices of the policy parameters (θ_π, θ_x) lead to large inflation or deflation and large welfare losses (see Section 5.3 for further discussion). We also examined the results when the policymaker uses a different loss function $\tilde{L} = Var(\pi) + \lambda_x Var(\tilde{x})$ where the loss depends on the volatility of filtered output gap estimates. Note that the equilibrium under learning is completely independent from the choice of loss function; thus, the learning equilibrium stays the same as in the above analysis. The choices of the policy parameters (θ_π, θ_x) which cause instability (large inflation volatility) in the baseline loss function continue to cause instability under the alternative loss function. Therefore, this alternative loss function leads to negligible changes to the (in)stability region associated with the OG estimation methods. However, the optimal policy now requires a stronger response to inflation and the optimal welfare loss becomes lower, because the volatility of the filtered output is generally lower than that of the true output gap (see e.g., Figures A.5 and A.6).

Alternative parameterization and modification of OG estimation methods. Wolf et al. (2026) argue for a smoothing parameter of 650 for the one-sided HP filter in real-time analysis, and Quast and Wolters (2022) propose a modification of the Hamilton filter based on averaging forecast horizons from one to three years. Appendix I reports the results for these two variants: both have substantially larger stability regions than their baseline counterparts (the HP filter with a smoothing parameter of 1600 and the original Hamilton filter), for all values of the interest rate smoothing parameter.

stability region for all detrending methods. Two main differences here are the presence of the ZLB and the explicit choice of detrending methods in our paper, issues which are not considered in Orphanides and Williams (2007).

[27] Another benchmark is the case when private agents are learning but the policymaker is assumed to know the true OG. The iso-loss contours in this scenario are shown on the bottom right plot of Figure 1. In this case, the stable region is small and smaller than almost all detrending methods.

Interpreting the ranking. Our framework favors methods which produce OG estimates with smaller amplitudes and volatility, and one might think that such methods merely attribute more variability to the trend estimate. This is true when the data generating process for aggregate output is exogenous, as typically assumed when econometricians evaluate estimation methods (e.g., Canova, 2025). In our framework, however, aggregate output is endogenously influenced by the choice of estimation method and interest rate decisions, so one method can produce more variability in *both* OG estimates and trend estimates relative to another, because it generates larger volatility of aggregate output itself.

Result 1 (Stability ranking)**.** *With the ZLB, the size of the stability region differs sharply across estimation methods: the CF bandpass, BN-KMW, HP and FOD filters have the largest stability regions, while deterministic detrending (linear, quadratic, DLS) has the smallest. The ranking is unchanged across degrees of interest rate inertia and across all extensions considered.*

5.2 Stability with the Taylor rule coefficients

The Taylor rule has been widely popular in the monetary policy literature since the publication of Taylor (1993) and is a frequently used benchmark in monetary models. We thus consider the optimal welfare achieved across OG estimation methods, conditional on the choice of the Taylor rule by the central bank. We provide a simulation of the model with the original Taylor rule coefficients, i.e., $\theta_\pi = 1.5$, $\theta_x = 0.125$ and $\theta_r = 0$. The model is simulated for 40,000 periods with all parameters and shocks identical across all detrending methods.

Table A.1 in Appendix F reports statistics associated with the simulation: the standard deviation of real-time estimates of OG, inflation, interest rate, inflation expectation and nominal interest rate expectation, the welfare loss and the proportion of simulation periods hitting the ZLB. Figure A.4, also in Appendix F, displays the real-time estimates of the OG with each method in the different panels.

As can be seen from the table and figure, some detrending methods intrinsically produce larger and more volatile OG estimates, such as the linear, quadratic, and DLS detrending. This will lead to a larger volatility in interest rates, more volatile beliefs of households, and a larger welfare loss. Other methods produce smaller and less volatile OG estimates, such as the bandpass filter and BN-KMW filter. Economies where these methods are adopted will have a smaller volatility in interest rates, household beliefs and a lower welfare loss.

Regardless of the detrending method chosen by the policymaker (except for DLS detrending), the Taylor rule coefficients robustly ensures stability of the economy (the rule is represented by

the red triangle in each panel of Figure 1).[28] Interestingly, the HP filter, BN-KMW filter, and CF bandpass filter yield the lowest welfare loss of 0.358.[29]

Result 2 (Robust policy prescription)**.** *Irrespective of the estimation method, a sufficiently large response to inflation combined with a small response to OG estimates—as in the original Taylor rule—keeps the economy stable and delivers good welfare performance.*

5.3 What drives instability and large welfare losses?

Why is there heterogeneity in the performance across different detrending methods in terms of the size of the stability regions? Simulations are provided to shed light on the (in)stability results. The two main messages are the following. First, generally, when the Taylor Principle is fulfilled, any detrending method which intrinsically produces large and volatile OG estimates is more prone to deflationary spirals and large welfare losses. Second, an insufficient response to inflation (violating the Taylor Principle) can lead to inflation spirals.

5.3.1 ZLB and large deflation

The combination of volatile OG estimates and the presence of the ZLB may lead to instability even when the Taylor Principle is fulfilled. To get some intuition, we compare two detrending methods, the BN-KMW filter and linear detrending, with very different stability regions (the former being much larger in size). The policy reaction coefficients are set to $\theta_\pi = 2.5$, $\theta_x = 0.7$, and $\theta_r = 0.6$. This choice is in the stability region associated with the BN-KMW filter but not for the linear detrending method. We shut down all shocks except feeding into the model a sequence of negative 3-standard-deviation shocks to the "IS" curve for 10 quarters, which could occur during a large recession.

The blue lines in Figure A.5 in Appendix F plot the trajectory of true OG, real-time OG estimates, inflation and nominal interest rate in the economy as well as the expectation of these variables for 200 periods under linear detrending; the red lines plot the same variables with the BN-KMW filter.[30]

[28] The sole exception is the DLS method which produces explosive dynamics and extremely large volatility of inflation, OG and welfare loss; hence it is not reported in Table A.1.

[29] We note that some of the methods which track the OG better do not necessarily yield smaller loss, for instance, the linear, quadratic and the Hamilton Filter are among the best in tracking the OG though not necessarily yielding the smallest loss.

[30] More precisely, the interest rate plotted is the deviation of the nominal interest rate from its long-run equilibrium rate, so the lower bound in the figure is $-(\pi^* + r^*)$, i.e., -0.015. Appendix F also plots the same simulations for 1000 periods and for two additional methods (quadratic detrending and HP filter).

Following the sequence of negative demand shocks, output and inflation start to decline, estimates of potential output are revised downward, and the linear detrending method produces a large negative OG estimate; nominal interest rates fall and reach the ZLB. After the shocks disappear, OG estimates and inflation start to recover and nominal rates rise before the economy has fully recovered. The linear detrending method then produces *positive* OG estimates while the true OG remains large and negative. The large overestimation of the OG requires significant hikes in policy rates, which cause inflation and the OG to fall again and put their paths in a downward spiral: the interest rate returns to the ZLB and stays there, and falling inflation raises real interest rates and their expectations, further reducing inflation and output. Put another way, volatile OG estimates cause nominal interest rates to be too high exactly when they should be low to help the economy recover from the recession.

With the BN-KMW filter, by contrast, nominal rates (on average) stay low at the ZLB as long as the economy is recovering from the large recession. Thereafter, with gradually rising inflation and OG, the interest rate eventually exits the ZLB and the economy recovers towards the steady state. This appears to be the key difference between detrending methods that lead to stability and those that do not: with unstable methods, over-estimation of the OG during the recovery phase keeps nominal rates too high, pushing the economy further into recession, with falling inflation and OG and interest rates stuck forever at the ZLB.[31]

5.3.2 Decomposing the role of ZLB constraint in driving instability

To isolate the role of the ZLB in driving the instability results, we calculate the size of the stability region, ignoring the ZLB constraint. Thus the policymaker is assumed to set the nominal interest rate according to the following policy rule:

$$i_t = \theta_r i_{t-1} + \theta_\pi \pi_{t-1} + \theta_x \widetilde{x}_{t-1} + e_{i,t}. \tag{5.1}$$

Panel B of Table 1 reports the size of stability region for the detrending methods without the ZLB. The iso-loss contours are displayed in Appendix G.1. The size of stability region for all detrending methods increase, compared with the ZLB case. In particular, with low policy inertia ($\theta_r = 0$ or $\theta_r = 0.2$), the size of stability region with every detrending method is about 90% and the differences in the size of stability region are no longer substantial.[32] With higher policy

[31] The mechanism also holds for other detrending methods. For instance, Figure A.8 of Appendix F shows that under the same parameterization and shocks, the economy with quadratic detrending (or HP filter) leads to instability (or stability).

[32] The exception is DLS detrending where the large and volatile OG estimates continue to play an important role in generating instability.

inertia ($\theta_r = 0.6$ or $\theta_r = 0.8$), the stability region of linear, quadratic and DLS detrending shrink considerably compared to the other methods.[33]

The results suggest that it is the combination of volatile OG estimates and the presence of the ZLB that makes some methods unstable: volatile OG estimates mean hitting the ZLB more frequently (or staying there longer), leading to deflation scenarios. In normal times when the ZLB does not bind, the choice of detrending method does not matter much for stability and welfare losses (except for the deterministic methods); in times when the ZLB binds (say during a crisis), the choice matters hugely. This is a clear warning for policymakers: they need to be especially careful about the choice of detrending method during severe recessions or in times of crisis.

5.3.3 High inflation when interest rates respond insufficiently to inflation

We demonstrate a different source of instability in our next simulation: an insufficient response of interest rates to inflation (violating the Taylor Principle). For illustration, we use the CF bandpass filter, which has a large stability region, and set $\theta_r = 0$, $\theta_\pi = 0.5$ and $\theta_x = 0.2$, a combination that violates the Taylor Principle and lies in the instability region of the bandpass filter. All shocks are included, as in the baseline case. Figure A.6 in Appendix F plots a typical trajectory: inflation rises (and the OG is positive and generally rising), and since the nominal interest rate does not react aggressively enough, real interest rates (and their expectations) are negative and falling, further fueling inflationary pressures and leading to divergence from the steady state. This type of explosive behavior is due more to the lack of adherence to the Taylor principle on the part of the policymaker than to the choice of the detrending method *per se*.

Result 3 (Mechanism and state dependence)**.** *Instability arises from the combination of volatile OG estimates and the ZLB: methods that intrinsically produce large and volatile estimates are prone to self-reinforcing deflation spirals when the ZLB binds, while an insufficient response to inflation generates inflation spirals. Without the ZLB, stability regions are large and broadly comparable across methods.*

[33] Iterating the rule backwards, higher policy inertia implies a higher weight on past OG estimates, which contributes to more instability with these methods.

6 Baseline results: optimal policy and reliability of detrending methods

Given the detrending method chosen by the policymaker, we first characterize optimal policies from the class of simple interest rules (2.7) and then the properties of revisions to OG estimates.

6.1 Optimal policy coefficients and welfare loss

Characterizing optimal policies within the class of simple interest rules has been a popular practice in the monetary policy literature, see, for example, Rotemberg and Woodford (1997). We conduct a similar exercise in spirit for different OG estimation methods in this sub-section. For each detrending method and across different degrees of interest rate inertia, Table 2 summarizes optimized policy coefficients, associated standard deviations of inflation and the OG, welfare losses and the proportion of simulation periods hitting the ZLB. Given the degree of inertia, the response to inflation in the optimal policy across all detrending methods is broadly similar in magnitude. When the degree of inertia is low (0 and 0.2), the optimal response to inflation is very high across all detrending methods (ranging from 5.3 to 7). When the degree of inertia is large (0.6 and 0.8), the optimal response to inflation is smaller across all detrending methods, though still more than one. With an inertia coefficient of 0.8, the responses are more of the kind observed in actual policy practice (1.5 – 1.7).

The optimized policy reaction coefficients to the OG estimates vary considerably across the detrending methods. In Panel B of Table 2, for the methods which intrinsically produce larger and more volatile OG estimates (e.g., polynomial detrending), the optimal response to OG estimates is typically small. In fact, it is zero (or less than 0.1) with degree of inertia 0.6 or 0.8 implying behavior compatible with a strict inflation targeting central bank. Holding other parameters constant, a larger reaction coefficient to OG estimates would increase the volatility of interest rate and beliefs of agents, which would lead to a larger inflation volatility and welfare loss. By contrast, Panel A suggests that for the methods which intrinsically produce smaller and less volatile OG estimates (e.g., the bandpass filter, BN-KMW filter), optimal policy dictates a larger response to OG estimates (mostly 0.2 – 0.6).

For all detrending methods, the policy coefficients reacting to inflation and the OG depend negatively on the degree of interest rate inertia. When the degree of inertia is higher (or lower), optimal policy requires a more muted (or aggressive) response to inflation and the OG. A higher degree of inertia means a larger dependence on the interest rate in the previous period which

in turn depends on past inflation and OGs: inflation and OGs are highly correlated over time. Thus, policy with a larger degree of inertia requires a less aggressive optimal response to current inflation and OG.

Regardless of the OG estimation method chosen by the policymaker, a sufficiently large response to inflation and small response to OG estimates, such as in the original Taylor rule, robustly deliver good welfare performance. Too weak a response to inflation will lead to positive inflation spirals and large welfare losses, as illustrated in Section 5.3. A large response to OG estimates, especially in the presence of the ZLB, will lead to deflation spirals and large welfare losses, especially for those methods which produce large and volatile OG estimates.

Given the degree of inertia, the difference in welfare loss associated with optimal policy across detrending methods is small. The optimal loss associated with the FOD filter, Hamilton filter, CF band-pass filter and BN-KMW filter are a bit lower than that of the remaining detrending methods. In particular, if the policymaker employs an interest rate rule with a high degree of inertia, FOD detrending outperforms all other methods by generating the smallest social welfare loss.

Despite the similarity in optimal loss levels across methods, the choice of method is far from irrelevant. As discussed in the Introduction, neither the policy coefficients adopted in practice nor the model parameters are known precisely; with a small stability region, departures from the optimized coefficients can easily tip the economy into instability, whereas a large stability region keeps such departures inexpensive. Methods with larger stability regions also tend to be more reliable (Section 6.2), which is why the stability region remains crucial in judging detrending methods.

The last column of Table 2 displays the proportion of simulation periods hitting the ZLB associated with the optimal policy. The chance of hitting the ZLB is a negative function of the degree of inertia θ_r. This is because a higher (or lower) degree of inertia reduces (increases) interest rate volatility. So interest rates will more (or less) likely fluctuate around the steady state and less (or more) likely hit the ZLB.

We note that a policymaker facing uncertainty about which estimation method to use would ideally derive robust policies that perform adequately across methods, perhaps using minimax or Bayesian model averaging approaches (Brock et al., 2003; Giannoni, 2002). Our finding that a sufficiently large response to inflation and a small response to OG estimates robustly delivers good performance across all methods provides a practical recommendation in this direction. However, a thorough analysis of these approaches is beyond the scope of the paper and left for

Table 2: OPTIMAL POLICY AND WELFARE LOSS

Method	Coeff		SD				Coeff		SD			
	θ_π	θ_x	σ_π	σ_x	Loss	ZLB(%)	θ_π	θ_x	σ_π	σ_x	Loss	ZLB(%)
Panel A: methods intrinsically producing smaller $\widetilde{x}_t$							**Panel B: methods intrinsically producing bigger $\widetilde{x}_t$**					
Christiano-Fitzgerald band-pass filter							Blanchard-Quah					
$\theta_r = 0$	7.0	0.9	0.477	4.855	0.298	52.2	$\theta_r = 0$	6.6	0.1	0.477	4.877	0.299	48.9
$\theta_r = 0.2$	5.7	0.8	0.476	4.860	0.297	47.8	$\theta_r = 0.2$	5.4	0.1	0.476	4.883	0.298	44.3
$\theta_r = 0.6$	2.9	0.5	0.475	4.899	0.298	35.4	$\theta_r = 0.6$	2.8	0.0	0.475	4.943	0.299	31.8
$\theta_r = 0.8$	1.7	0.4	0.475	4.988	0.301	30.6	$\theta_r = 0.8$	1.5	0.0	0.475	5.055	0.303	23.8
Beveridge-Nelson (KMW)							Beveridge-Nelson (bivariate)						
$\theta_r = 0$	6.7	0.5	0.477	4.854	0.298	50.2	$\theta_r = 0$	6.6	0.1	0.477	4.877	0.299	48.9
$\theta_r = 0.2$	5.4	0.4	0.476	4.860	0.298	45.1	$\theta_r = 0.2$	5.4	0.1	0.476	4.888	0.298	44.3
$\theta_r = 0.6$	2.9	0.4	0.475	4.884	0.297	35.4	$\theta_r = 0.6$	2.8	0.0	0.475	4.943	0.299	31.8
$\theta_r = 0.8$	1.8	0.3	0.475	4.996	0.301	31.5	$\theta_r = 0.8$	1.5	0.0	0.475	5.055	0.303	23.8
HP filter							Hamilton						
$\theta_r = 0$	6.9	0.3	0.477	4.877	0.299	50.8	$\theta_r = 0$	6.7	0.2	0.477	4.846	0.298	50.1
$\theta_r = 0.2$	5.3	0.2	0.476	4.878	0.298	43.9	$\theta_r = 0.2$	5.6	0.2	0.476	4.852	0.297	46.7
$\theta_r = 0.6$	2.9	0.3	0.475	4.919	0.298	35.0	$\theta_r = 0.6$	3.1	0.2	0.475	4.876	0.297	39.0
$\theta_r = 0.8$	1.6	0.2	0.475	5.016	0.302	27.6	$\theta_r = 0.8$	1.7	0.1	0.475	4.994	0.301	29.7
First-order differencing							Quadratic trend						
$\theta_r = 0$	6.9	0.6	0.477	4.841	0.298	53.2	$\theta_r = 0$	6.8	0.1	0.477	4.876	0.299	50.3
$\theta_r = 0.2$	5.6	0.6	0.477	4.836	0.297	48.9	$\theta_r = 0.2$	5.5	0.1	0.476	4.884	0.298	45.4
$\theta_r = 0.6$	3.1	0.6	0.477	4.840	0.297	41.1	$\theta_r = 0.6$	2.8	0.0	0.475	4.943	0.299	31.8
$\theta_r = 0.8$	2.1	0.6	0.476	4.884	0.298	41.1	$\theta_r = 0.8$	1.5	0.0	0.475	5.055	0.303	23.8
Beveridge-Nelson (1981)							Linear trend						
$\theta_r = 0$	6.6	0.1	0.477	4.872	0.299	49.3	$\theta_r = 0$	6.6	0.1	0.477	4.841	0.298	49.5
$\theta_r = 0.2$	5.4	0.2	0.476	4.873	0.298	45.4	$\theta_r = 0.2$	5.3	0.1	0.476	4.839	0.297	44.7
$\theta_r = 0.6$	2.9	0.4	0.476	4.884	0.298	37.2	$\theta_r = 0.6$	3.1	0.1	0.475	4.886	0.298	37.7
$\theta_r = 0.8$	1.7	0.3	0.476	4.972	0.301	32.1	$\theta_r = 0.8$	1.5	0.0	0.475	5.055	0.303	23.8
							Discounted Least Squares						
							$\theta_r = 0$	6.6	0.0	0.477	4.877	0.299	48.9
							$\theta_r = 0.2$	5.4	0.0	0.476	4.887	0.298	44.2
							$\theta_r = 0.6$	2.8	0.0	0.475	4.943	0.299	31.7
							$\theta_r = 0.8$	1.5	0.0	0.475	5.055	0.303	23.9
							$\widetilde{x}_{t-1} = x_{t-1}$						
							$\theta_r = 0$	4.6	0.2	0.472	4.451	0.282	41.8
							$\theta_r = 0.2$	4.1	0.2	0.472	4.455	0.282	42.6
							$\theta_r = 0.6$	2.1	0.1	0.472	4.504	0.283	32.4
							$\theta_r = 0.8$	1.9	0.1	0.476	4.706	0.293	44.9

Notes: The table reports the optimized policy coefficients (θ_π, θ_x), the associated standard deviations and welfare loss for each detrending method under different degrees of interest rate inertia (θ_r). Panel A groups methods that intrinsically produce smaller OG estimates; Panel B groups those producing larger OG estimates. The interest rate rule is $i_t = \max\{\theta_r i_{t-1} + \theta_\pi \pi_{t-1} + \theta_x \widetilde{x}_{t-1} + e_{i,t}, -(\pi^* + r^*)\}$. SD: standard deviation of inflation (σ_π) and the output gap (σ_x). ZLB (%): percentage of simulation periods hitting the zero lower bound. Standard deviations and welfare losses are multiplied by 100 and 10,000, respectively, to facilitate readability.

future research.

Assuming the policymaker knows the true OG under households' learning and sets interest rates according to the true OG, the bottom right plots of Figure 1 and Figures A.1-A.3 in Appendix E show that the size of the stability region is very small. This is because the true OG is large and volatile, which can easily lead to a large volatility in agents' beliefs and explosive dynamics. Table 2 suggests that the optimal response to the OG is small. The loss under the optimal policy in this case is lower than the optimal loss achieved under the policymaker's learning with any of the detrending methods (though higher than that under full-information RE). Thus, knowledge of the true OG entails a trade off in that the lower welfare loss comes at the cost of a smaller stability region.

Result 4 (Optimal policy)**.** *Given the degree of inertia, the optimal response to inflation is similar across methods, whereas the optimal response to OG estimates varies inversely with the size and volatility of the estimates a method produces; differences in minimized welfare loss across methods are small.*

6.2 Properties of revisions to output gap estimates

A large literature evaluates the reliability of (or properties of revisions to) OG estimates from detrending methods, e.g. Orphanides and van Norden (2002), Kamber et al. (2018), Edge and Rudd (2016) and Champagne et al. (2018). A reliability measure is calculated as the average of absolute revisions to OG estimates over a number of years where the revisions are computed as ex post estimates minus real-time estimates of OG. A detrending method is considered more reliable if the associated reliability measure is smaller (or the average size of revisions to the OG estimate is smaller). The revisions are made when new data observations arrive.[34]

The literature typically uses the same historical output data to compute the reliability measure for each detrending method. This evaluation practice does not take into account the interaction between detrending methods, policy decisions and macroeconomic outcomes: a policymaker adopting a different method obtains different OG estimates, sets different interest rates and generates different macroeconomic outcomes, which are then used to estimate the OG in the following period. We re-examine reliability while explicitly recognizing this dynamic interaction.

In the same spirit, we define the model-produced revision in the OG estimate as $\widetilde{x}_{t-k|t} - \widetilde{x}_{t-k|t-k}$, where $\widetilde{x}_{t-k|t}$ is the ex post estimate for period $(t-k)$ based on detrending results in

[34]The paper omits issues related to revisions of output data, which is a smaller factor contributing to revisions of OG estimates, as shown in Orphanides and van Norden (2002).

period t and $\widetilde{x}_{t-k|t-k}$ the real-time estimate at period $t-k$; we examine 10-year revisions ($k=40$) and report the average of their absolute values. The model with the optimized policy coefficients with no inertia is simulated for 40,000 periods for each detrending method. Table 3 reports five statistics of the revisions: the mean absolute revision, the standard deviation, minimum and maximum values, and the first-order serial correlation.[35]

The BN-KMW filter is the most reliable method, producing the smallest revisions to OG estimates: the mean absolute revision and the standard deviation, maximum and minimum of the revisions are 0.16%, 0.20%, 0.70% and -0.83%, respectively. The filter was introduced by Kamber et al. (2018) precisely to mitigate the unreliability of many trend-cycle decomposition methods brought out by Orphanides and van Norden (2002), by explicitly fixing a low signal-to-noise ratio. Our findings demonstrate the robustness of this result even when account is taken of the endogeneity between detrending and the data generating process.

Table 3: PROPERTIES OF REVISIONS TO OUTPUT GAP ESTIMATES

Method	*Mean*	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald bandpass	0.63	0.80	3.11	-3.47	0.938
Beveridge-Nelson (KMW, 2016)	0.16	0.20	0.70	-0.83	0.973
HP filter	1.02	1.28	5.18	-5.69	0.974
Beveridge-Nelson (1981) filter	1.33	1.51	6.53	-8.25	0.703
Blanchard-Quah filter	1.48	2.17	9.56	-11.19	0.905
Beveridge-Nelson (bivariate)	1.49	2.41	12.78	-12.37	0.881
Hamilton filter	0.56	0.72	4.54	-3.97	0.875
Quadratic trend	2.19	2.75	9.81	-10.15	0.997
Linear trend	2.70	3.36	10.24	-11.71	0.999

Notes: The table reports the properties of revisions to real-time output gap estimates under the optimal policy for each detrending method. Revisions are calculated as the ex post OG estimate made 10 years later ($\widetilde{x}_{t-k|t}$, $k=40$) minus the real-time estimate ($\widetilde{x}_{t-k|t-k}$). *Mean*: average of absolute OG estimate revisions; *SD*: standard deviation of revisions; *Max* and *Min*: maximum and minimum revision values; *AR*: first-order serial correlation of revisions. DLS and FOD are excluded because their OG estimates are invariant to the rolling-window sample.

The CF bandpass filter and Hamilton filter are less reliable than the BN-KMW filter but more reliable than the remaining detrending methods. Linear and quadratic detrending are the least reliable and generate very large revisions: the mean and standard deviation of their revisions are more than twice those generated by the HP filter. The multivariate methods, i.e., the bivariate BN filter and the BQ filter, are no more reliable than many univariate methods, such as the BN-KMW, bandpass, Hamilton and HP filters. Consistent with the literature, the revisions are generally

[35]The FOD and DLS filters are excluded from the table: by definition the FOD filter is not associated with any revision, while the DLS filter is unreliable and produces very large revisions.

highly autocorrelated; see the last column. These findings are in line with Orphanides and van Norden (2002), corroborated here by our richer model specific analysis.

Result 5 (Reliability)**.** *The BN-KMW filter produces the smallest revisions to OG estimates. Stability and reliability are closely associated: methods with large stability regions are also reliable, despite the data generating process being endogenous to the method choice.*[36]

7 Extensions and robustness

This section briefly summarizes two sets of exercises: extensions to alternative monetary policy regimes and robustness checks with respect to our modeling assumptions. At the outset, we stress that our main results are robust throughout. The ranking of detrending methods in terms of the size of the stability region does not change in any scenario. Moreover, the optimal response to OG estimates with methods which intrinsically produce larger and more volatile (or smaller and less volatile) OG estimates is smaller (or bigger). A sufficiently large response to inflation and small response to OG estimates robustly delivers good welfare performance, irrespective of the detrending method. Detailed model specifications and full numerical results are collected in the appendix; here we outline each exercise and its main lessons.

Alternative monetary policy regimes. We consider a 4% inflation target (Appendix J), a -1% lower bound on nominal interest rates (Appendix G), a dual mandate that places a higher relative weight on OG volatility in the loss function (Yellen, 2012; Debortoli et al., 2019; Appendix K), and average inflation targeting (AIT) as recently adopted by the US Fed (Appendix L). A higher inflation target, a negative lower bound and AIT all alleviate the ZLB problem: interest rates hit the bound less often, which reduces the volatility of inflation and the OG and hence the volatility of households' and the policymaker's beliefs, dampening the mechanism that generates self-reinforcing deflation spirals. As a result, the stability region expands with every detrending method, welfare under the optimal policy improves, and it is generally optimal to respond somewhat more strongly to OG estimates. Under the dual mandate, only the loss function changes; the stability regions and the properties of OG revisions are identical to the baseline, while the optimal policy shifts toward a larger response to OG estimates and a smaller response to inflation.

Robustness. Whether the underlying output trend is deterministic or stochastic is a long-standing open question (Nelson and Plosser, 1982; Gordon, 1997), and policymakers in real time

[36]There are some exceptions, e.g., the Hamilton filter has a relatively small stability region but a relatively high degree of reliability.

may be uncertain about a potential acceleration or slowdown in long-run growth rates (e.g., Holston et al., 2017). Appendix M therefore repeats the analysis when the underlying trend is deterministic and when it contains shocks to both its level and growth rate, and characterizes the robust optimal interest rate rule against such uncertainty by using a Min-Max algorithm, following Orphanides and Williams (2007). The Min-Max monetary policy strategies coincide with the optimal policy in the case with both level and growth shocks to the trend, the worst-case scenario. Appendix N considers slower and faster learning by households; the size of the stability region depends negatively on the gain parameters. Appendix O adopts the model specification and parameterization of Orphanides and Williams (2007, 2008) and endows households with a correctly specified forecasting model that nests the RE solution (as in, for instance, Eusepi and Preston, 2011; Kuang, 2014), in contrast to the simple, underparameterized rules in the baseline case. Finally, Appendix P considers an infinite horizon learning formulation motivated in Preston (2005, 2006), where agents' consumption and price setting decisions depend on their forecasts into the entire infinite future. In all of these scenarios, the main results—including the ranking of detrending methods and the character of optimal policy—remain unchanged.

8 Conclusion

We develop a New Keynesian model in which the policymaker has imperfect knowledge of the economy, forms real-time estimates of the unobserved output gap (OG) with one of eleven recommended estimation methods, and commits to a simple interest rate rule subject to the zero lower bound (ZLB). We numerically characterize the conditions under which monetary policy leads to self-reinforcing deflation or inflation spirals and large welfare losses; the choice of estimation method plays a crucial role in these conditions. The BN-KMW, CF bandpass, HP and FOD filters have the largest stability regions, whereas the deterministic methods (linear, quadratic, DLS) have small ones—a ranking that is surprisingly robust across alternative policy regimes (a higher inflation target, a dual mandate, average inflation targeting, a negative lower bound), trend specifications and belief specifications.

Beyond optimal policy, we evaluate the methods by the reliability criterion of Orphanides and van Norden (2002). A number of their findings survive the dynamic interaction between estimation, policy and outcomes in our model, and stability and reliability turn out to be closely associated: methods with large stability regions are also quite reliable.

While we deliberately choose a simple three-equation New Keynesian model as a benchmark to analyze the welfare and monetary policy implications of eleven OG estimation methods,

it would be worthwhile to extend the analysis along a number of dimensions: enriching the model with capital, capacity utilization and financial frictions, which can significantly affect OG estimates and their policy implications (Justiniano and Primiceri, 2008; Slobodyan and Wouters, 2012); adding learning about the natural rate of interest or unemployment by the policymaker and analyzing the role of policy errors (Orphanides and Williams, 2007); deriving welfare criteria from first principles under bounded rationality (Benchimol and Bounader, 2023; Benchimol, 2024); and considering robust policies that perform adequately across estimation methods, following the literature on model uncertainty and robustness (Brock et al., 2003; Hansen and Sargent, 2008; Giannoni, 2002). More broadly, the framework developed in the paper can be used to evaluate other existing or new estimation methods, for any country and using any preferred economic model (e.g. models managed by central banks). We leave these for future work.

References

Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *Economic Journal*, 117(520):603–636.

Adam, K., Kuang, P., and Marcet, A. (2012). House price booms and the current account. *NBER Macroeconomics Annual*, 26:77–122.

Armona, L., Fuster, A., and Zafar, B. (2019). Home price expectations and behavior: Evidence from a randomized information experiment. *Review of Economic Studies*, 86:1371–1410.

Ascari, G. and Sbordone, A. (2014). The macroeconomics of trend inflation. *Journal of Economic Literature*, 52(3):679–739.

Benchimol, J. (2024). Central bank objectives, monetary policy rules, and limited information. *Journal of Macroeconomics*, 80:1–21.

Benchimol, J. and Bounader, L. (2023). Optimal monetary policy under bounded rationality. *Journal of Financial Stability*, 89:1–15.

Bernanke, B. and Woodford, M. (1997). Inflation forecasts and monetary policy. *Journal of Money, Credit and Banking*, 29(4):653–684. Part 2: Dynamic Effects of Monetary Policy.

Beveridge, S. and Nelson, C. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, 7:151–174.

Binder, C., Kuang, P., and Tang, L. (2026). Central bank communication and house price expectations. *Journal of the European Economic Association*, 24(2):769–799.

Blanchard, O., Dell'Ariccia, G., and Mauro, P. (2010). Rethinking macroeconomic policy. *Journal of Money, Credit and Banking*, 42:199–215.

Blanchard, O. and Galí, J. (2007). Real wage rigidities and the new keynesian model. *Journal of Money, Credit and Banking*, 39(1):35–65.

Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79(4):655–673.

Brainard, L. (2019). Federal reserve review of monetary policy strategy, tools, and communications: Some preliminary views. Remarks at the Presentation of the 2019 William F. Butler Award, New York Association for Business Economics, New York.

Brock, W. A., Durlauf, S. N., and West, K. D. (2003). Policy evaluation in uncertain economic environments. *Brookings Papers on Economic Activity*, 34(1):235–322.

Budianto, F., Nakata, T., and Schmidt, S. (2023). Average inflation targeting and the interest rate lower bound. *European Economic Review*, 152:104384.

Bullard, J. and Keating, J. (1995). The long-run relationship between inflation and output in postwar economies. *Journal of Monetary Economics*, 36(3):477–496.

Bullard, J. and Mitra, K. (2002). Learning about monetary policy rules. *Journal of Monetary Economics*, 49(6):1105–1129.

Bullard, J. and Mitra, K. (2007). Determinacy, learnability, and monetary policy inertia. *Journal of Money, Credit and Banking*, 39(5):1177–1212.

Calvo, G. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Canova, F. (2025). FAQ: How do I estimate the output gap? *Economic Journal*, 135(665):59–80.

Canova, F. and Sala, L. (2009). Back to square one: Identification issues in DSGE models. *Journal of Monetary Economics*, 56(4):431–449.

Champagne, J., Poulin-Bellisle, G., and Sekkel, R. (2018). The real-time properties of the Bank of Canada's staff output gap estimates. *Journal of Money, Credit and Banking*, 50(6):1167–1188.

Christiano, L. and Fitzgerald, T. (2003). The band pass filter. *International Economic Review*, 44(2):435–465.

Christiano, L. J. and Gust, C. (1999). Comment on "Robustness of simple monetary policy rules under model uncertainty" by Andrew Levin, Volker Wieland and John C. Williams. In Taylor, J. B., editor, *Monetary Policy Rules*, pages 299–316. University of Chicago Press, Chicago.

Clarida, R., Galí, J., and Gertler, M. (1999). The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37(4):1661–1707.

Coibion, O., Gorodnichenko, Y., and Ulate, M. (2018). The cyclical sensitivity in estimates of potential output. *Brookings Papers on Economic Activity*, 2018(1):343–441.

Coibion, O., Gorodnichenko, Y., and Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):1537–1584.

Coibion, O., Gorodnichenko, Y., and Wieland, J. (2012). The optimal inflation rate in new keynesian models: Should central banks raise their inflation targets in light of the zero lower bound? *Review of Economic Studies*, 79:1371–1406.

Congressional Budget Office (2020). An update to the economic outlook: 2020 to 2030. Technical report, Government Printing Office, Washington, DC.

Davig, T. and Doh, T. (2014). Monetary policy regime shifts and inflation persistence. *Review of Economics and Statistics*, 96(5):862–875.

Debortoli, D., Kim, J., Lindé, J., and Nunes, R. (2019). Designing a simple loss function for central banks: Does a dual mandate make sense? *The Economic Journal*, 129(621):2010–2038.

Di Pace, F., Mitra, K., and Zhang, S. (2021). Adaptive learning and labor market dynamics. *Journal of Money, Credit and Banking*, 53(2-3):441–475.

Edge, R. and Rudd, J. (2016). Real-time properties of the Federal Reserve's output gap. *Review of Economics and Statistics*, 98(4):785–791.

Eusepi, S. and Preston, B. (2011). Expectations, learning, and business cycle fluctuations. *American Economic Review*, 101(6):2844–2872.

Eusepi, S. and Preston, B. (2018). The science of monetary policy: An imperfect knowledge perspective. *Journal of Economic Literature*, 56(1):3–59.

Fuster, A., Laibson, D., and Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives*, 24(4):67–84.

Gabaix, X. (2020). A behavioral new keynesian model. *American Economic Review*, 110(8):2271–2327.

Giannoni, M. P. (2002). Does model uncertainty justify caution? robust optimal monetary policy in a forward-looking model. *Macroeconomic Dynamics*, 6(1):111–144.

Giannoni, M. P. and Woodford, M. (2004). Optimal inflation targeting rules. In Bernanke, B. and Woodford, M., editors, *The Inflation-Targeting Debate*, pages 93–162. University of Chicago Press, Chicago.

Gordon, S. (1997). Stochastic trends, deterministic trends, and business cycle turning points. *Journal of Applied Econometrics*, 12:411–434.

Hamilton, J. (2018). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics*, 100(5):831–843.

Hansen, L. P. and Sargent, T. J. (2008). *Robustness*. Princeton University Press, Princeton.

Hodrick, R. and Prescott, E. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Holston, K., Laubach, T., and Williams, J. (2017). Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics*, 108:S59–S75.

Ireland, P. (2004). Technology shocks in the new keynesian model. *Review of Economics and Statistics*, 86(4):923–936.

Justiniano, A. and Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3):604–641.

Kamber, G., Morley, J., and Wong, B. (2018). Intuitive and reliable estimates of the output gap from a Beveridge-Nelson filter. *Review of Economics and Statistics*, 100(3):550–566.

Knotek, E. (2007). How useful is Okun's law? *Federal Reserve Bank of Kansas City Economic Review*, pages 73–103.

Kuang, P. (2014). A model of housing and credit cycles with imperfect market knowledge. *European Economic Review*, 70:419–437.

Kuang, P. and Mitra, K. (2016). Long-run growth uncertainty. *Journal of Monetary Economics*, 79:67–80.

Kuang, P. and Mitra, K. (2025). Potential output pessimism and austerity in the European Union. *Journal of Money, Credit and Banking*, 57(7):1871–1905.

Kuang, P., Tang, L., Zhang, R., and Zhang, T. (2022). Forecast disagreement about long-run macroeconomic relationships. *Journal of Economic Behavior and Organization*, 200:371–387.

Kuang, P., Tang, L., Zhang, R., and Zhang, T. (2025). Are survey stock price forecasts anchored by fundamental forecasts? A long-run perspective. *Economic Theory*, 79(2):657–685.

Laubach, T. and Williams, J. (2003). Measuring the natural rate of interest. *Review of Economics and Statistics*, 85(4):1063–1070.

Levin, A., Wieland, V., and Williams, J. (2003). The performance of forecast-based monetary policy rules under model uncertainty. *American Economic Review*, 93:622–645.

Luo, S. and Startz, R. (2014). Is it one break or ongoing permanent shocks that explains US real GDP? *Journal of Monetary Economics*, 66:155–163.

Marcet, A. and Sargent, T. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–368.

Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics*, 54:2065–2082.

Morley, J., Nelson, C., and Zivot, E. (2003). Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, 85(2):235–243.

Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10:139–162.

Okun, A. (1962). Potential output: Its measurement and significance. In *American Statistical Association 1962 Proceedings of the Business and Economic Section*, Washington, DC. American Statistical Association.

Orphanides, A. (2002). Monetary-policy rules and the great inflation. *American Economic Review*, 92(2):115–120.

Orphanides, A. and van Norden, S. (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics*, 84(4):569–583.

Orphanides, A. and Williams, J. (2007). Robust monetary policy with imperfect knowledge. *Journal of Monetary Economics*, 54:1406–1435.

Orphanides, A. and Williams, J. (2008). Learning, expectations formation, and the pitfalls of optimal control monetary policy. *Journal of Monetary Economics*, 55:S80–S96.

Perron, P. and Wada, T. (2009). Let's take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics*, 56(6):749–765.

Preston, B. (2005). Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking*, 1(2):81–126.

Preston, B. (2006). Adaptive learning, forecast-based instrument rules and monetary policy. *Journal of Monetary Economics*, 53(3):507–535.

Quast, J. and Wolters, M. (2022). Reliable real-time output gap estimates based on a modified Hamilton filter. *Journal of Business and Economic Statistics*, 40(1):152–168.

Rotemberg, J. and Woodford, M. (1996). Real-business-cycle models and the forecastable movements in output, hours, and consumption. *American Economic Review*, 86:71–89.

Rotemberg, J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual*, 12:297–346.

Rotemberg, J. and Woodford, M. (1999). Interest rate rules in an estimated sticky price model. In Taylor, J. B., editor, *Monetary Policy Rules*, pages 57–126. University of Chicago Press, Chicago.

Slobodyan, S. and Wouters, R. (2012). Learning in an estimated medium-scale DSGE model. *Journal of Economic Dynamics and Control*, 36(1):26–46.

Smets, F. (2003). Maintaining price stability: How long is the medium term? *Journal of Monetary Economics*, 50(6):1293–1309.

Stock, J. and Watson, W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44:293–335.

Summers, L. (2014). U.S. economic prospects: Secular stagnation, hysteresis, and the zero lower bound. *Business Economics*, 49(2):65–73.

Svensson, L. E. O. (2010). Inflation targeting. In Friedman, B. M. and Woodford, M., editors, *Handbook of Monetary Economics*, volume 3, chapter 22. Elsevier.

Svensson, L. E. O. (2011). Practical monetary policy: Examples from Sweden and the United States. *Brookings Papers on Economic Activity*, pages 289–332.

Svensson, L. E. O. (2020). Monetary policy strategies for the Federal Reserve. *International Journal of Central Banking*, 16:133–193.

Swanson, E. and Williams, J. (2014). Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review*, 104(10):3154–3185.

Taylor, J. (1999). A historical analysis of monetary policy rules. In Taylor, J. B., editor, *Monetary Policy Rules*, pages 319–341. University of Chicago Press, Chicago.

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214.

Ulate, M. (2021). Going negative at the zero lower bound: The effects of negative nominal interest rates. *American Economic Review*, 111(1):1–40.

van Norden, S. (2002). Filtering for current analysis. Bank of Canada Working Paper 2002-28, Bank of Canada.

Williams, J. (2006). Monetary policy in a low inflation economy with learning. Working paper, Federal Reserve Bank of San Francisco.

Wolf, E., Mokinski, F., and Schüler, Y. (2026). On adjusting the one-sided Hodrick-Prescott filter. *Journal of Money, Credit and Banking*, 58:919–931.

Woodford, M. (1999). Optimal monetary policy inertia. *The Manchester School*, 67(s1):1–35.

Woodford, M. (2002). Imperfect common knowledge and the effects of monetary policy. In Aghion, P., Frydman, R., Stiglitz, J., and Woodford, M., editors, *Knowledge, Information and Expectations in Modern Macroeconomics*. Princeton University Press, Princeton.

Woodford, M. (2003). *Interest and Prices*. Princeton University Press, Princeton.

Yellen, J. (2012). The economic outlook and monetary policy. Speech by Vice Chair Yellen at the Money Marketeers of New York University, New York.

Online Appendix

A Output gap estimation methods: an overview

This appendix provides an overview of the output gap estimation methods studied in the paper.

Deterministic trends. The first category of detrending methods assumes that the trend component in (log) output is well approximated and represented by a straight-forward deterministic function of time. The linear detrending is one of the simplest methods that fits the trend component by a straight line and the quadratic detrending additionally attempts to capture the potential curvature. The linear or quadratic detrending method removes a linear or quadratic time trend using a sample size of T and the resulting time series is our estimates of output gaps. For instance, by OLS estimation, the quadratic detrending runs the regression

$$y_t = a + bt + ct^2 + \epsilon_t, \tag{A.1}$$

to estimate the trend, which is defined as $y^{tr} = a + bt + ct^2$ and the residual is the cyclical component. By construction, all fluctuations in the observations are concluded into the cyclical component from the linear or quadratic detrending, including business cycles, high- and low-frequency variations.

Lansing (2002) assumes that the Fed constructs a series of output gap estimate using an algorithm of Discounted Least Squares (DLS) which minimizes the sum of squared difference between a linear trend and a series of past observations of output $\{d^{T-t} \cdot y_t\}_{t=1}^{T}$ that are "discounted" at a constant rate, $0 \leq d \leq 1$. This additionally introduced parameter is analogous to the gain parameter (constant gain = 1-d) in a recursive learning model, where an agent cares more about recent observations and their attention to past observations gradually fades away. We also examine the case where the policymaker in our model adopts this detrending method by assuming a relatively large $d = 0.99$.

First-order differencing. The second type of method is the first-order differencing (FOD) filter, which assumes the trend equals the level of output from the previous period, $y_t^{tr} = y_{t-1}$, and extracts output gap as the difference:

$$\widetilde{x}_t = y_t - y_{t-1}. \tag{A.2}$$

Due to the unobservable nature of true output gaps, central banks sometimes input other output-related measure in lieu of output gap in monetary policy rules, such as output growth. FOD filter is equivalent to using real-time output growth in the policy. Canova (2025) finds that given a

simulated series of output, FOD outperforms many elaborate filters in estimating output gaps.

Hodrick-Prescott filter. The third type of method is the commonly used Hodrick-Prescott trend filter of Hodrick and Prescott (1997) (HP filter). Unlike the first category of methods, the HP trend is assumed to be stochastic, which also nests smoothing splines. We implement a one-sided derivation of HP filter introduced by Stock and Watson (1999) for real-time analysis and later adopted by Orphanides and van Norden (2002) to examine the reliability of real-time output gap estimation. It decomposes a series under the specification of an unobserved-components form. In addition to equation (3.5), the trend component is generally assumed to be

$$\Delta^d y_t^{tr} = \eta_t, \tag{A.3}$$

where Δ^d is the dth-order difference operator ($\Delta = 1 - L$, L is the lag operator) for $d > 0$. The cyclical component y_t^c is identified as a white noise that is mutually uncorrelated with the white noise η_t. Following Orphanides and van Norden (2002), we focus on the case $d = 2$. Note that equation (A.3) can then be alternatively written as

$$y_t^{tr} = 2y_{t-1}^{tr} - y_{t-2}^{tr} + \eta_t, \tag{A.4}$$

where the trend is an unobserved state variable (see e.g. Harvey and Jaeger, 1993; King and Rebelo, 1993). Particularly, the ratio of variance of η_t and of y_t^c is fixed as $q = \frac{\sigma_\eta^2}{\sigma_{y^c}^2}$, which is the reciprocal of the smoothing parameter in the two-sided HP filter. We set q to $\frac{1}{1600}$ in this paper. The policymaker identifies the filtered series as the estimated trend (potential output) and the residual as the cyclical (output gap) estimate.

Hamilton filter. The fourth method is a regression-based method proposed by Hamilton (2018). It decomposes a time series into a stochastic trend and cyclical component by regressing the following model using Ordinary Least Square (OLS):

$$y_{t+i} = \beta_0 + \sum_{l=0}^{L} \beta_{l+1} y_{t-l} + e_{t+i}, \tag{A.5}$$

where i denotes the ith lead in quarter and l denotes the lth lag in quarter from time t. The trend at time $t+i$ is defined as the fitted value from the regression,

$$\hat{y}_{t+i}^{tr} = \hat{y}_{t+i} = \hat{\beta}_0 + \sum_{l=0}^{L} \hat{\beta}_{l+1} y_{t-l}, \tag{A.6}$$

and the cyclical component is the deviation of the actual value from the estimated trend

$$\widetilde{x}_{t+i} = y_{t+i} - \hat{y}_{t+i}^{tr}. \tag{A.7}$$

Following Hamilton (2018), we assume the policymaker implements the Hamilton filter by choosing the lead of $i = 8$ quarters and the lag of $l = 4$ quarters for our quarterly model.

Band-pass filter. The fifth category of decomposition is the band-pass filter (Baxter and King, 1999; Christiano and Fitzgerald, 2003; Harvey and Trimbur, 2003). The band-pass filter decomposes the output series following (3.5) and isolates the cyclical component y_t^c whose period of oscillation falls within a chosen range (w^l, w^u). This feature directly enables the concept of business cycle to be operational and definable. Particularly, y_t^c outside the frequency in the interval $\{(-a,-b) \cup (a,b)\}$, $0 < a < b < \pi$ is wiped out and similarly y_t^{tr} has power only in the complement set of the interval $\{(-a,-b) \cup (a,b)\}$ within $(-\pi,\pi)$, where $a = \frac{2\pi}{w^u}$ and $b = \frac{2\pi}{w^l}$. The mechanism eliminates both high- and low-frequency variability in the data sample by filtering out the fluctuations in a chosen range (w^l, w^u). A two-sided ideal band-pass filter

$$B(L) = \sum_{j=-\infty}^{\infty} B_j L^j,$$

where

$$B_j = \frac{\sin(jb) - \sin(ja)}{\pi j}, \; j \neq 0$$
$$B_0 = \frac{b-a}{\pi}, a = \frac{2\pi}{w^u}, b = \frac{2\pi}{w^l}$$

filters in the time domain but optimally weights for each observation in the frequency domain. It typically requires an infinitely long data sample. Christiano and Fitzgerald (2003) constructs a version of approximated asymmetric ideal band-pass filter with a finite sample:

$$\hat{y}_t = \hat{B}_t(L) y_t = \sum_{j=-n_1}^{n_2} \hat{B}_{t,j} y_{t+j}, \tag{A.8}$$

where

$$\hat{B}_{t,j} = \operatorname{argmin} E\{(y_t^c - \hat{y}_t)^2\}, \tag{A.9}$$

where $n_1 = t-1$ and $n_2 = T-t$ for a sample of size T (so that the filter uses all available observations) and $y_t^c = B(L) y_t$ is the ideal band-pass filtered series. Following the recommendation of Christiano and Fitzgerald (CF henceforth, 2003), the policymaker is assumed to choose $w^l =$ 1.5 years (6 quarters) and $w^u = 8$ years (32 quarters). The results are not very sensitive to the selection of w^l in the range $[1,3]$ and w^u in the range $[6,10]$.

Univariate Beveridge-Nelson decompositions. Beveridge and Nelson (BN henceforth, 1981) estimate the trend for a unit-root times series as the long-run predictable component (the forecast as horizon approaches infinity) from which the average of deterministic drift is removed.

Its ARIMA representation has the following form:

$$\Delta y_t = \mu + A(L)\epsilon_t, \tag{A.10}$$

where $A(L) = 1 + A_1 L + A_2 L^2 + ...$, $\epsilon_t \sim IIN(0, \sigma_\epsilon^2)$ and μ is the drift parameter (the drift parameter can be either estimated as the sample mean or the long-run mean of series Δy_t). Beveridge and Nelson (1981) demonstrate that the minimum square error forecasts of $y_{t+\tau}$ based on information available at t is $E_t y_{t+\tau} = y_t + E_t(\Delta y_{t+1} + \Delta y_{t+2} + ... + \Delta y_{t+\tau}) = y_t + \sum_{j=1}^{\tau} E_t \Delta y_{t+j}$. Using equation (A.10), $E_t y_{t+\tau} = y_t + \tau\mu + (\sum_{j=1}^{\tau} A_j)\epsilon_t + (\sum_{j=2}^{\tau+1} A_j)\epsilon_{t-1} + ...$ Therefore,

$$y_t^{tr} = \lim_{\tau\to\infty}(E_t y_{t+\tau} - \tau\mu) = \lim_{\tau\to\infty}[y_t + (\sum_{j=1}^{\tau} A_j)\epsilon_t + (\sum_{j=2}^{\tau+1} A_j)\epsilon_{t-1} + ...] = y_{t-1}^{tr} + \mu + (1 + \sum_{j=1}^{\infty} A_j)\epsilon_t. \tag{A.11}$$

The trend follows a random walk with a constant drift. Finally, the cycle $\widetilde{x}_t = y_t - y_t^{tr}$. The feature of Beveridge and Nelson (1981) decomposition is that it imposes no restriction on the structure of the components and, in fact, uses a forecast-based method to identify the trend. Following Morley et al. (2003), we implement the decomposition with an AR(p) forecasting model

$$\Delta y_t = c + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + e_t, \tag{A.12}$$

where $e_t \sim IIN(0, \sigma_e^2)$. The policymaker is assumed to implement the Beveridge and Nelson (1981) filter with *p=2*.[37]

Kamber et al. (KMW henceforth, 2018) introduces a modified BN filter to improve the reliability of output gap estimates by generating relatively small output gap revision across time. The BN filter with a low signal-to-noise ratio transforms the AR(p) model of equation (A.12) into a Dickey-Fuller representation:

$$\Delta \widetilde{y}_t = \rho \Delta \widetilde{y}_{t-1} + \sum_{j=1}^{p-1} \phi_j^* \Delta^2 \widetilde{y}_{t-j} + e_t, \tag{A.13}$$

where $\rho = \phi_1 + \phi_2 + ... + \phi_p = 1 - \phi(1)$ and $\phi_j^* = -(\phi_{j+1} + ... + \phi_p)$. The signal-to-noise ratio is defined as $\delta = (1-\rho)^{-2}$ for the AR(p) model. Therefore, δ can be fixed by selecting ρ. Kamber et al. (2018) shows that when imposing a low signal-to-noise ratio, the modified BN filter produces small output gap revision. Following the recommendation of KMW, we assume the policymaker implements the BN-KMW filter with a high-order *AR(p)* model, *i.e. p=12,* and a low signal-to-

[37] KMW show that the Beveridge and Nelson (1981) filter with *p=1* implies a high signal-to-noise ratio and the resultant output gap estimate fails to match U.S. expansions and recessions determined by the National Bureau of Economic Research (NBER). Therefore, we choose *p=2* for the Beveridge and Nelson (1981) filter. We also experimented with the case *p=1*, which produces output gap estimate with small amplitude and low persistence.

noise ratio, $\rho = 0.05$. Their results (and our results below) are not sensitive to the selection of a higher-order *AR(p)* and $\rho \in [0.01, 0.08]$.

Multivariate Beveridge-Nelson decomposition. Following Kamber et al. (2016), we assume the policymaker implements the multivariate Beveridge-Nelson decomposition, by incorporating unemployment as the additional variable in the decomposition. In reference to Equation (A.10), under the multivariate setting,

$$\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \mu_{y1} \\ \mu_{y2} \end{bmatrix} + \sum_{i=1}^{k} \begin{bmatrix} \phi_{i,11} & \phi_{i,12} \\ \phi_{i,21} & \phi_{i,22} \end{bmatrix} \begin{bmatrix} \Delta y_{t-i} \\ u_{t-i} \end{bmatrix} + \begin{bmatrix} e_t^1 \\ e_t^2 \end{bmatrix},$$

where the level of output y_t is $I(1)$, unemployment u_t is $I(0)$ and e_t^1 is uncorrelated with e_t^2. We implement this with the number of lags $k = 8$.

Blanchard-Quah Structural VAR. Coibion et al. (2018) suggest that structural VARs (sVAR), such as Blanchard-Quah (BQ henceforth, 1989), produce output gap estimates that outperform filtering techniques not responding to the demand shocks. Blanchard-Quah sVAR has the same form as multivariate Beveridge-Nelson decompisition, but gives an economic interpretation to the shocks of the model, i.e. a demand shock and supply shock. The VAR system can be written as

$$\begin{bmatrix} 1 & \phi_{0,12} \\ \phi_{0,21} & 1 \end{bmatrix} \begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \mu_{y1} \\ \mu_{y2} \end{bmatrix} + \sum_{i=1}^{k} \begin{bmatrix} \phi_{i,11} & \phi_{i,12} \\ \phi_{i,21} & \phi_{i,22} \end{bmatrix} \begin{bmatrix} \Delta y_{t-i} \\ u_{t-i} \end{bmatrix} + \begin{bmatrix} e_t^s \\ e_t^d \end{bmatrix}$$

where e_t^d is the demand shock and uncorrelated with the supply shock e_t^s. The identifying restriction is that demand shocks have no effect on the level of GNP in the long run. Following Blanchard and Quah (1989), we set the number of lags $k = 8$.

General discussion of methods. A wide variety of OG estimation techniques are considered in the paper: these eleven techniques range from those most heavily used to more recently developed ones. These methods include purely statistical ones as well as ones with a theoretical foundation. We view an OG estimation method as having a theoretical foundation (and not purely statistical) if it makes use of economic theories or if the output process underlying the OG estimation method corresponds to that in economic models. The majority of methods (at least six) in this sense have a theoretical foundation. We explain this in more detail below.

For instance, the paper analyzes the OG estimation method (the Blanchard-Quah filter) proposed by Coibion et al. (2018). This filter estimates a structural VAR and recovers the component of aggregate output which responds only to the permanent shock as the estimate of potential output using the Blanchard and Quah identification strategy. Clearly, this method is not

purely statistical and utilizes economic theories to produce the OG estimate. In addition, three other methods, the celebrated Beveridge-Nelson decomposition (Beveridge and Nelson, 1981), a modification of Beveridge-Nelson filter (Kamber et al., 2018), and multivariate Beveridge-Nelson decomposition methods may not be viewed as purely statistical methods. The Introduction of Beveridge and Nelson (1981) provides strong theoretical motivations (dating back at least to the theories of Milton Friedman and John Muth) which led to the development of their decomposition method. The output process underlying these methods are consistent with neoclassical stochastic growth models with a balanced growth path (see, e.g., the Introduction of Rotemberg and Woodford, 1996, for a nice exposition on this point). Furthermore, the Hamilton filter too may not be viewed as purely statistical as the underlying output process is flexible and can correspond to the output process in stochastic and deterministic growth and DSGE models. In fact, Hamilton (2018) views this feature of his filter as a major advantage over the HP filter. Similarly, even the deterministic detrending methods (e.g., linear detrending) may not be viewed as purely statistical because the underlying output process corresponds to aggregate output in deterministic growth models (which dominated the modelling of economic growth for a long period before the development of stochastic growth models). Moreover, there have been debates on whether the underlying output trend is stochastic or deterministic, which remains an open question (e.g., Nelson and Plosser, 1982; Gordon, 1997).

While we consider estimation methods with a theoretical foundation, we do not want to restrict ourselves to these methods alone since even simple or statistical filtering methods are heavily used by policymakers all over the world and serve as crucial inputs in their monetary, fiscal, or macroprudential policy decisions. Thus, we feel incorporating a broad range of filters in our analysis is a strength of the paper.

B Calculation of output gap forecast

This appendix illustrates how the household's output gap forecasts are derived. Since the household is learning about output growth and their perceived law of motion of output growth follows equation (3.3), where their subjective output growth forecast at time t (based on $t-1$ information) is denoted by $\gamma_{t-1}^{\Delta y}$. The two-quarter ahead output gap forecast x_{t+1}^e is calculated as:

$$\begin{aligned} x_{t+1}^e &= y_{t+1}^e - (y_{t+1}^p)^e \\ &= 2\gamma_{t-1}^{\Delta y} + y_{t-1} - (2g + y_{t-1}^p) \qquad \text{(A.1)} \\ &= 2\gamma_{t-1}^{\Delta y} - 2g + x_{t-1}. \qquad \text{(A.2)} \end{aligned}$$

Similarly, the (N+1)-quarter ahead output gap forecast x^e_{t+N} is calculated as

$$\begin{aligned} x^e_{t+N} &= y^e_{t+N} - (y^p_{t+N})^e \\ &= (N+1)\gamma^{\Delta y}_{t-1} + y_{t-1} - \left[(N+1)g + y^p_{t-1}\right] \qquad \text{(A.3)} \\ &= (N+1)(\gamma^{\Delta y}_{t-1} - g) + x_{t-1}. \qquad \text{(A.4)} \end{aligned}$$

C Estimating the IS curve

We obtain the following OLS estimates for the unemployment gap model using the sample spanning from 1971:Q1 to 2019:Q4.[38]

$$u_t = 0.95 u^e_{t+1} + 0.02 u_{t-1} + \underset{(0.013)}{0.037}(i^e_t - \pi^e_{t+1}) + e_{u,t}, \qquad DW = 1.22, \qquad \text{(A.1)}$$

$$e_{u,t} = \underset{(0.068)}{0.390} e_{u,t-1} + v_{u,t}, \qquad \sigma_v = 0.29, \qquad \text{(A.2)}$$

$$\pi_t = 0.99\pi^e_{t+1} - \underset{(0.091)}{0.215}(u^e_t - u^*_t) + \widetilde{e}_{\pi,t}, \qquad DW = 1.20, \qquad \text{(A.3)}$$

$$\widetilde{e}_{\pi,t} = \underset{(0.069)}{0.360}\widetilde{e}_{\pi,t-1} + \widetilde{v}_{\pi,t}, \qquad \widetilde{\sigma}_\pi = 1.62, \qquad \text{(A.4)}$$

where u_t denotes unemployment and u^*_t represents the non-accelerating inflation rate of unemployment (NAIRU). Standard errors are reported in parentheses under each of estimated regression coefficients.[39] Autocorrelation in residuals is detected in the model by the Durbin-Watson test (DW = 1.22 and 1.20 for the IS and the Phillips curve, respectively) and the autocorrelation coefficients of the residuals are also estimated. A small negative correlation (about -6%) exists between two residuals.

D Calibrating the learning gains

The constant gain parameters, k^π, k^i and $k^{\Delta y}$ in households' belief updating are calibrated by minimizing the distance between the forecasts generated by adaptive learning and the corresponding survey forecasts. We feed actual inflation (measured as the percentage change in GDP or GNP price index), 3-month Treasury bill rates, and real GDP growth individually into the belief updating rule (3.4) to generate three series of forecasts. The in-sample period covers 1971:Q1 to 2019:Q4 and initial beliefs are set equal to the 5-year average values of inflation, 3-month Treasury bill rates, and real GDP growth, respectively, prior to the start of the in-sample period.

[38] Forecasts of 3-month Treasury bill rate in the first few years were not surveyed by SPF. Following Orphanides and Williams (2008), these forecasts are recovered by using the slope of the term structure under the expectations hypothesis.

[39] The estimation of (A.1) - (A.4) is performed with annualized variables (except for unemployment rate u_t and NAIRU u^*_t which are quarterly).

The gain parameters that result in the minimum mean square errors (MSEs) with regards to SPF forecasts are selected. MSEs are computed as:

$$MSE(w_j) = \frac{1}{t_1 - t_0} \sum_{t=t_0}^{t_1} (w_{j,t}^{model} - \widetilde{w}_t^{SPF})^2,$$

where $w_{j,t}^{model}$ is the one-quarter ahead forecast of variable w, under the jth gain parameter in the grid at the tth quarter, and $\widetilde{w}_t^{SPF}$ is the corresponding one-quarter ahead SPF forecast at the tth quarter. t_0 is 1971:Q1 and t_1 is 2019:Q4. This minimization exercise yields the gain parameters, $k^{\pi} = 0.008$, $k^{i} = 0.004$ and $k^{\Delta y} = 0.003$.

E Analysis under alternative degrees of interest rate inertia

This Appendix illustrates iso-loss contours under alternative degrees of interest rate inertia for the baseline model when the policymaker adopts one of the eleven methods to estimate the output gap. Figures A.1-A.3 plot the iso-loss contours for three alternative degrees of inertia, $\theta_r = 0.2$, 0.6 and 0.8. The details of the results are reported in Section 5. From the figures, it is clear that the size of the stability regions varies a lot across the different methods. In general, the stability regions associated with linear, quadratic, and discounted least squares detrending are rather small, whereas those associated with the BN-KMW, CF band-pass, HP, FOD and the BN filters are relatively large.

F Response to a series of negative shocks

This appendix collects simulation results underlying Sections 5.2 and 5.3. Table A.1 reports statistics from the simulation of the model with the original Taylor rule coefficients, and Figure A.4 displays the corresponding real-time estimates of the OG with each method.

Figure A.5 plots the trajectory of the economy under linear detrending (blue) versus the BN-KMW filter (red) following a sequence of negative demand shocks over 200 periods. Figure A.6 plots an inflationary spiral under the CF bandpass filter when the interest rate responds insufficiently to inflation. Figure A.7 plots the response of the economy with the two detrending methods (BN-KMW filter and linear detrending) to the sequence of negative demand shocks over a long period (1000 periods), and Figure A.8 plots the response of the economy for several detrending methods over a shorter horizon.

Figure A.1: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS ($\theta_r = 0.2$)

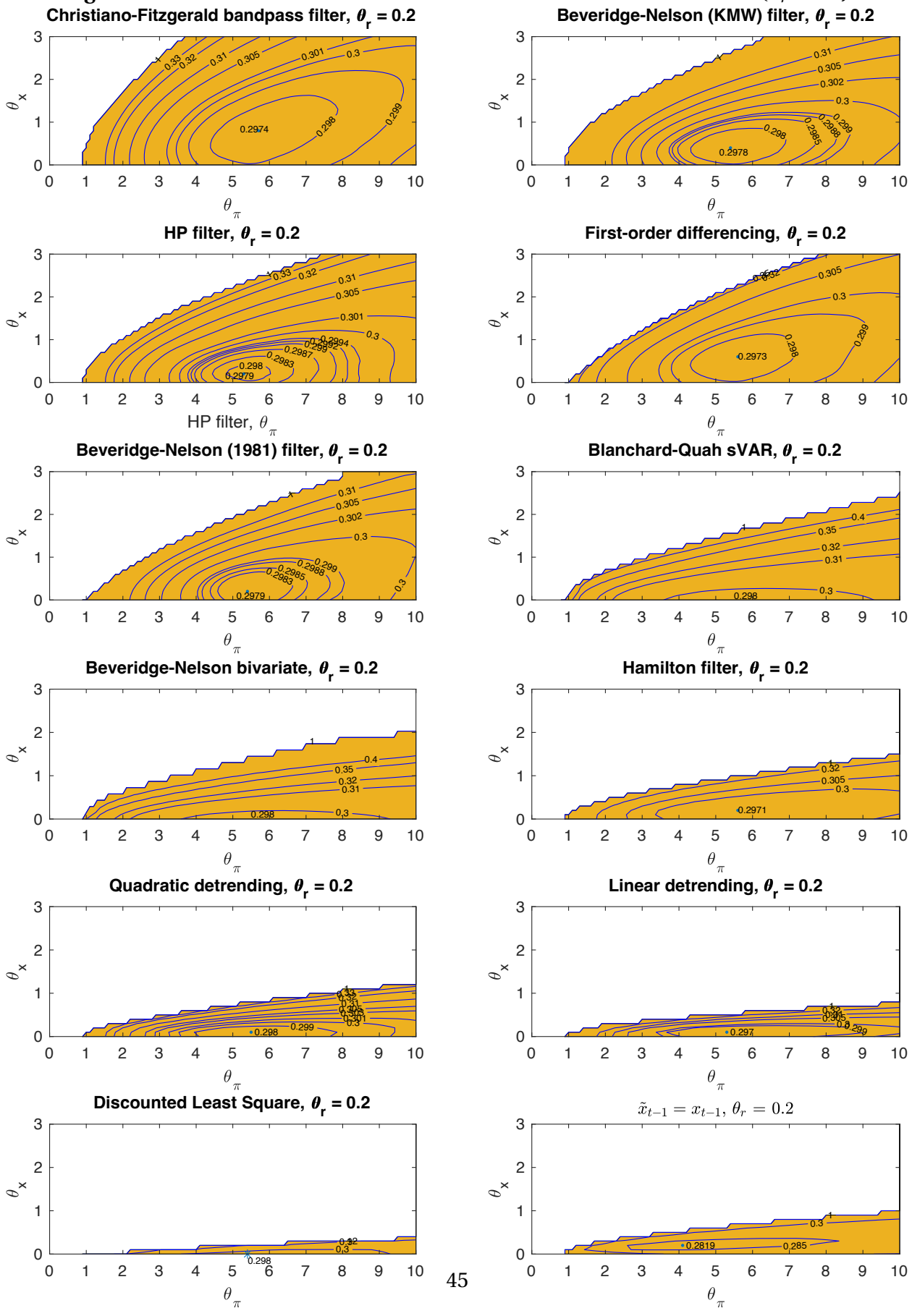


Figure A.2: ISO-LOSS CONTOURS WITH DETRENDING METHODS ($\theta_r = 0.6$)

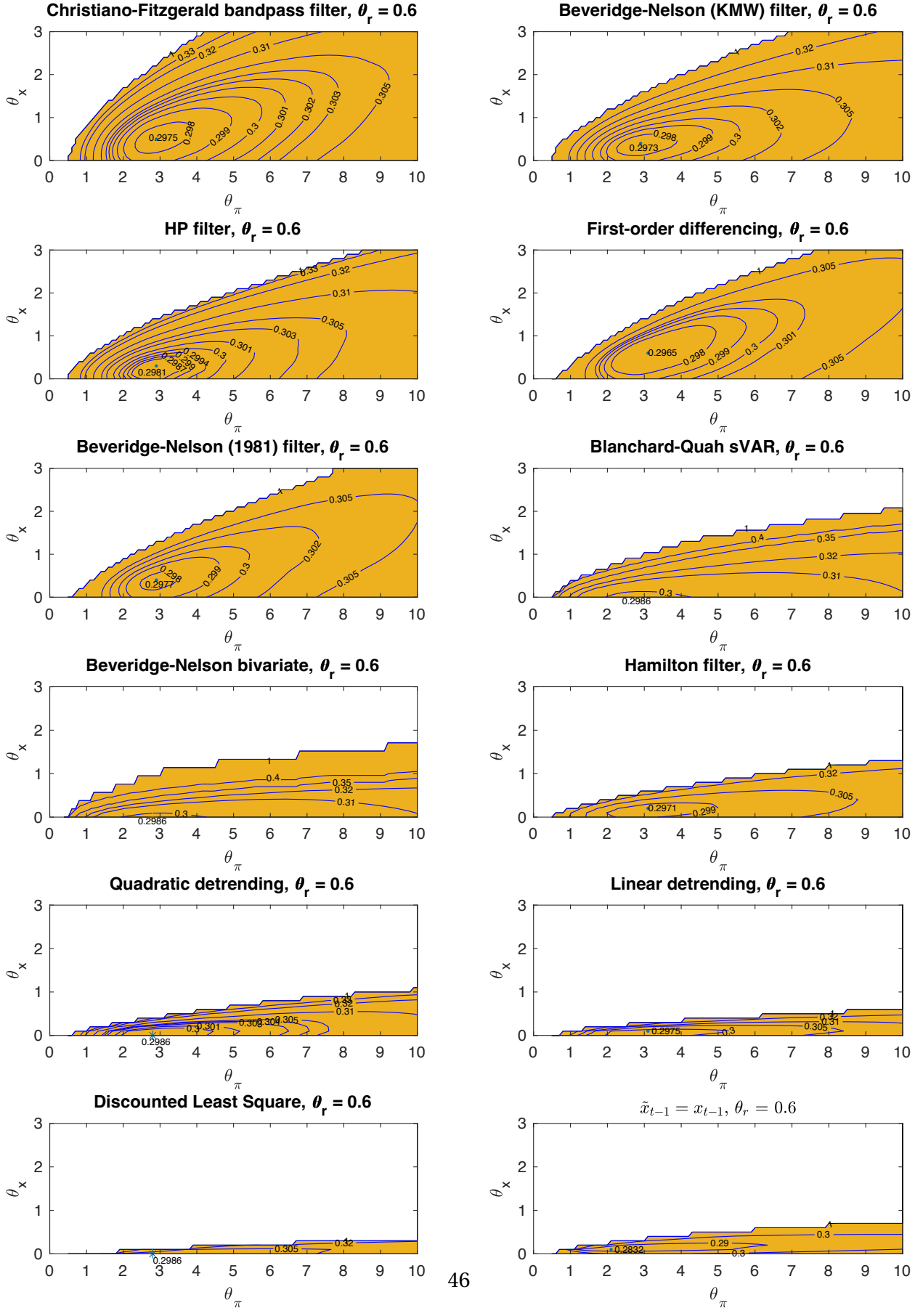


Figure A.3: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS ($\theta_r = 0.8$)

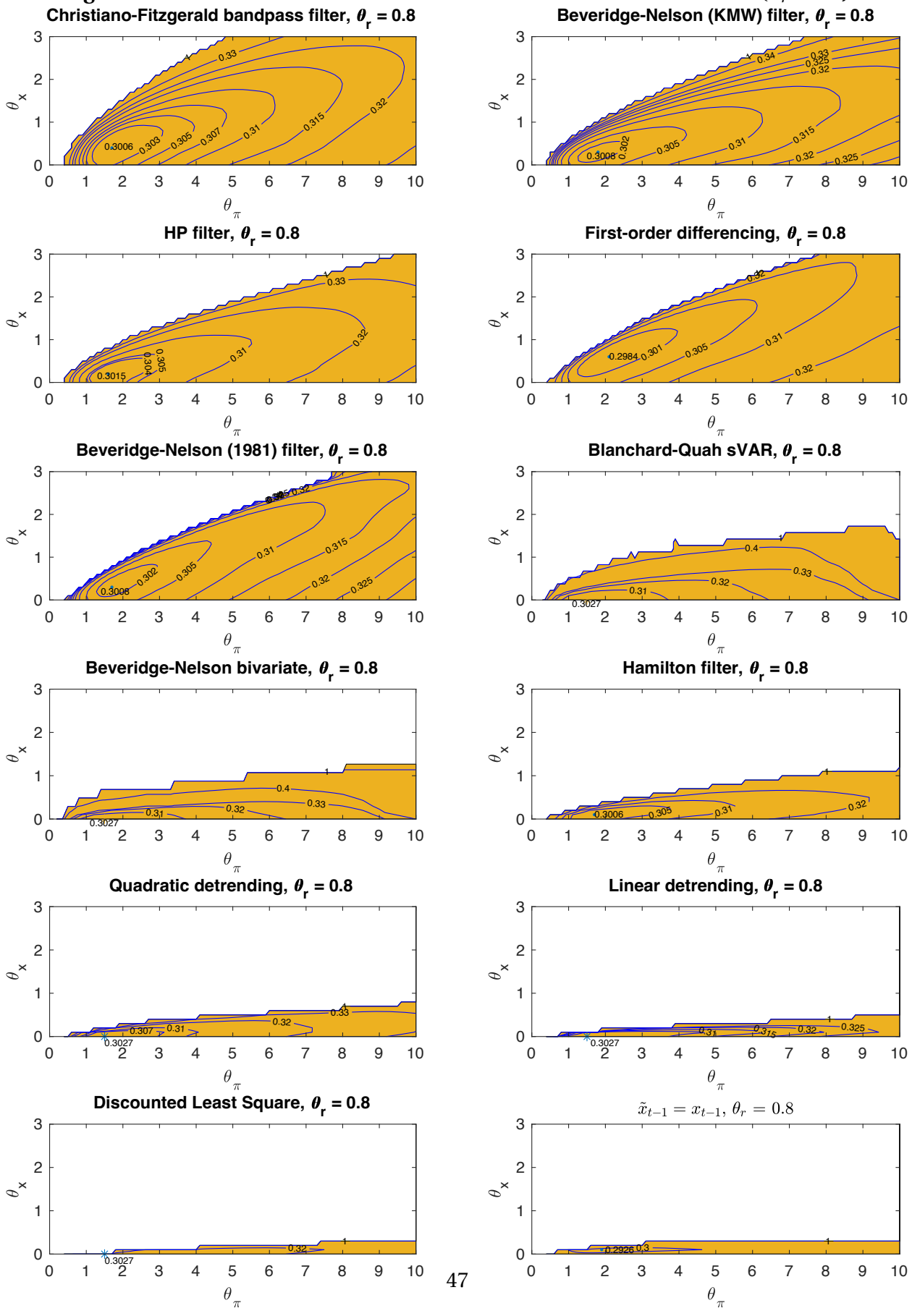


Table A.1: MODEL STATISTICS WITH TAYLOR RULE COEFFICIENTS

Method	$\sigma_{\tilde{x}}$	σ_π	σ_r	$\sigma_{E\pi}$	σ_{Er}	Loss	ZLB (%)
Christiano-Fitzgerald bandpass	0.789	0.537	0.900	0.283	0.370	0.358	6.635
Beveridge-Nelson (KMW)	1.086	0.537	0.902	0.284	0.368	0.358	6.773
HP filter	1.333	0.538	0.906	0.285	0.369	0.358	6.908
First-order differencing	0.802	0.545	0.879	0.299	0.376	0.366	10.285
Beveridge-Nelson (1981)	1.615	0.548	0.892	0.304	0.380	0.369	7.250
Blanchard-Quah	2.207	0.543	0.933	0.295	0.379	0.366	8.073
Beveridge-Nelson (bivariate)	2.218	0.548	0.934	0.304	0.388	0.372	8.270
Hamilton	2.783	0.543	0.945	0.296	0.377	0.363	8.375
Quadratic trend	3.123	0.541	0.952	0.291	0.373	0.362	8.765
Linear trend	3.787	0.549	0.984	0.307	0.383	0.368	10.360
$\widetilde{x}_{t-1} = x_{t-1}$	4.319	0.506	1.019	0.222	0.311	0.312	12.210

Notes: The table reports model-simulated statistics under the original Taylor rule coefficients ($\theta_\pi = 1.5$, $\theta_x = 0.5$, $\theta_r = 0$) for each detrending method. $\sigma_{\tilde{x}}$: standard deviation of the real-time OG estimate; σ_π: standard deviation of inflation; σ_r: standard deviation of the nominal interest rate; $\sigma_{E\pi}$: standard deviation of inflation expectations; σ_{Er}: standard deviation of interest rate expectations. Loss: welfare loss computed from equation (8). ZLB (%): percentage of simulation periods in which the zero lower bound binds. All standard deviations are multiplied by 100.

Figure A.4: REAL-TIME ESTIMATES OF THE OG FROM DIFFERENT DETRENDING METHODS WITH THE TAYLOR RULE COEFFICIENTS

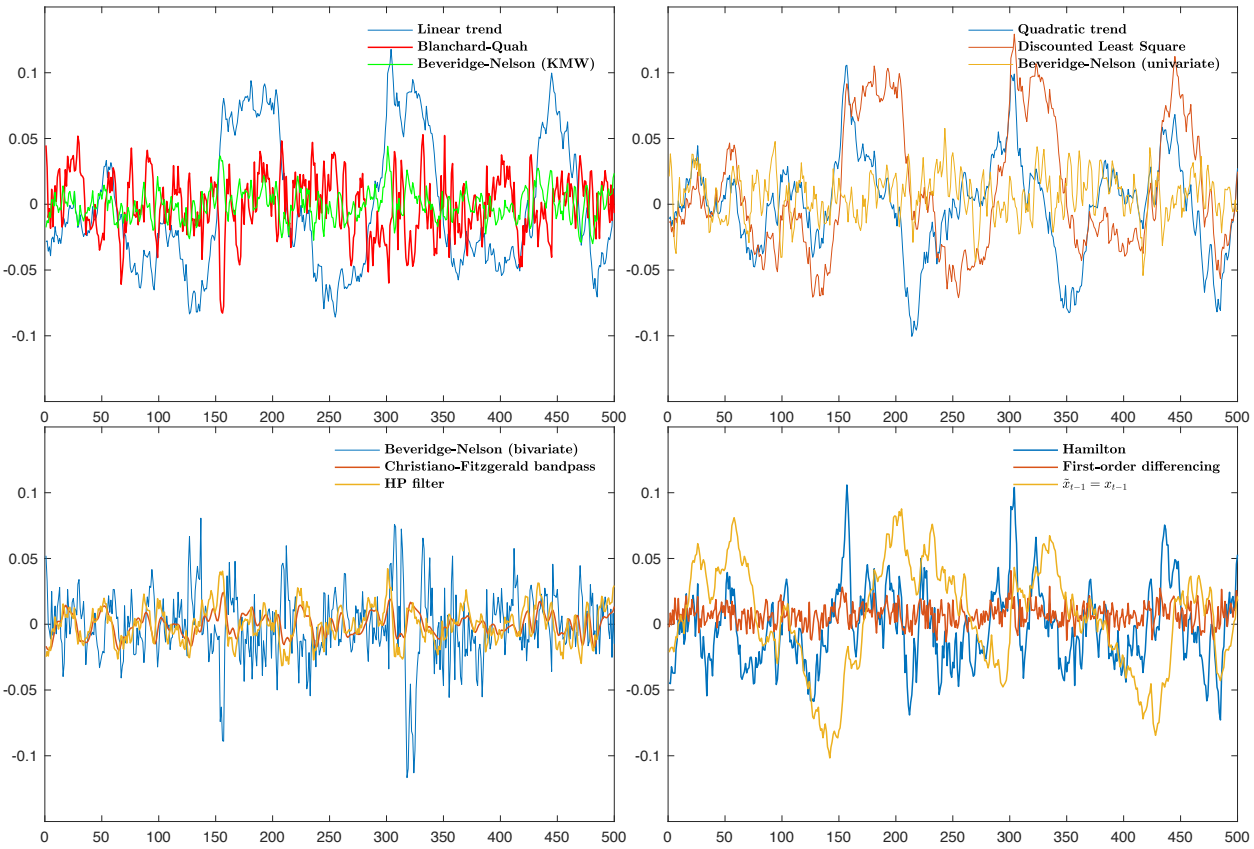


Figure A.5: THE ECONOMY WITH LINEAR DETRENDING (A DEFLATIONARY SPIRAL) VS BN-KMW FILTER FOLLOWING NEGATIVE DEMAND SHOCKS

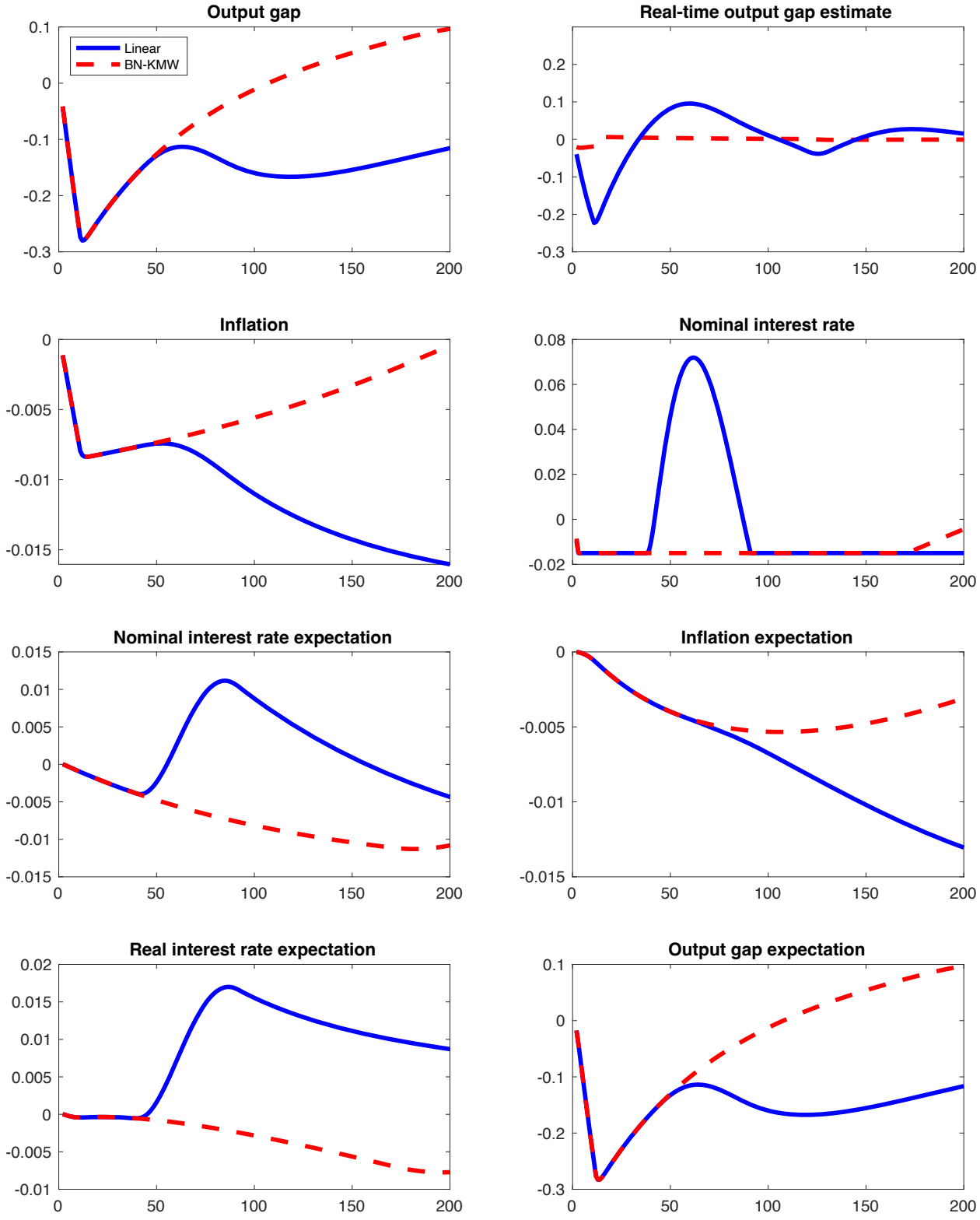


Figure A.6: INFLATIONARY SPIRAL WITH THE BANDPASS FILTER

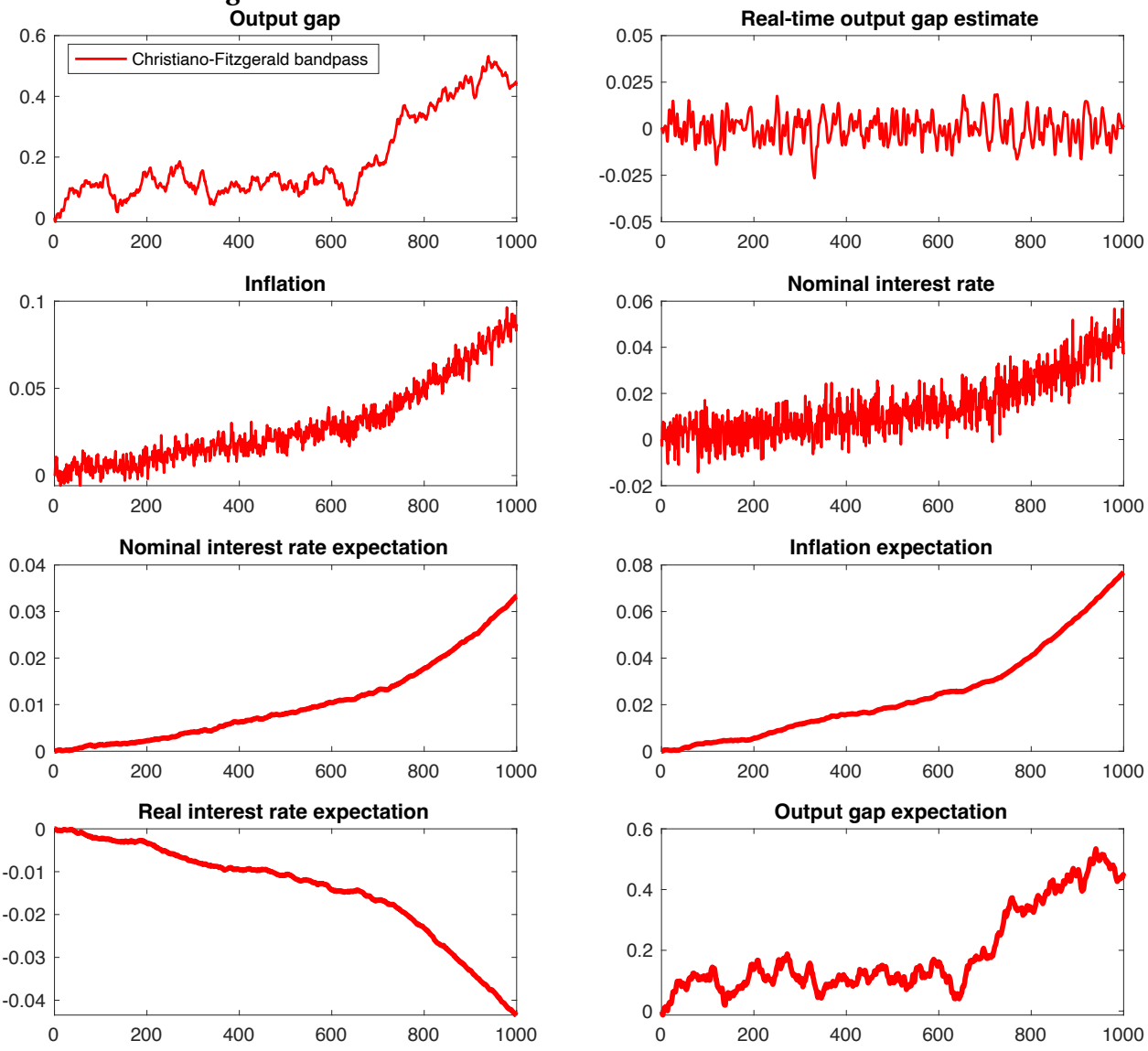


Figure A.7: RESPONSE TO A SEQUENCE OF NEGATIVE DEMAND SHOCKS

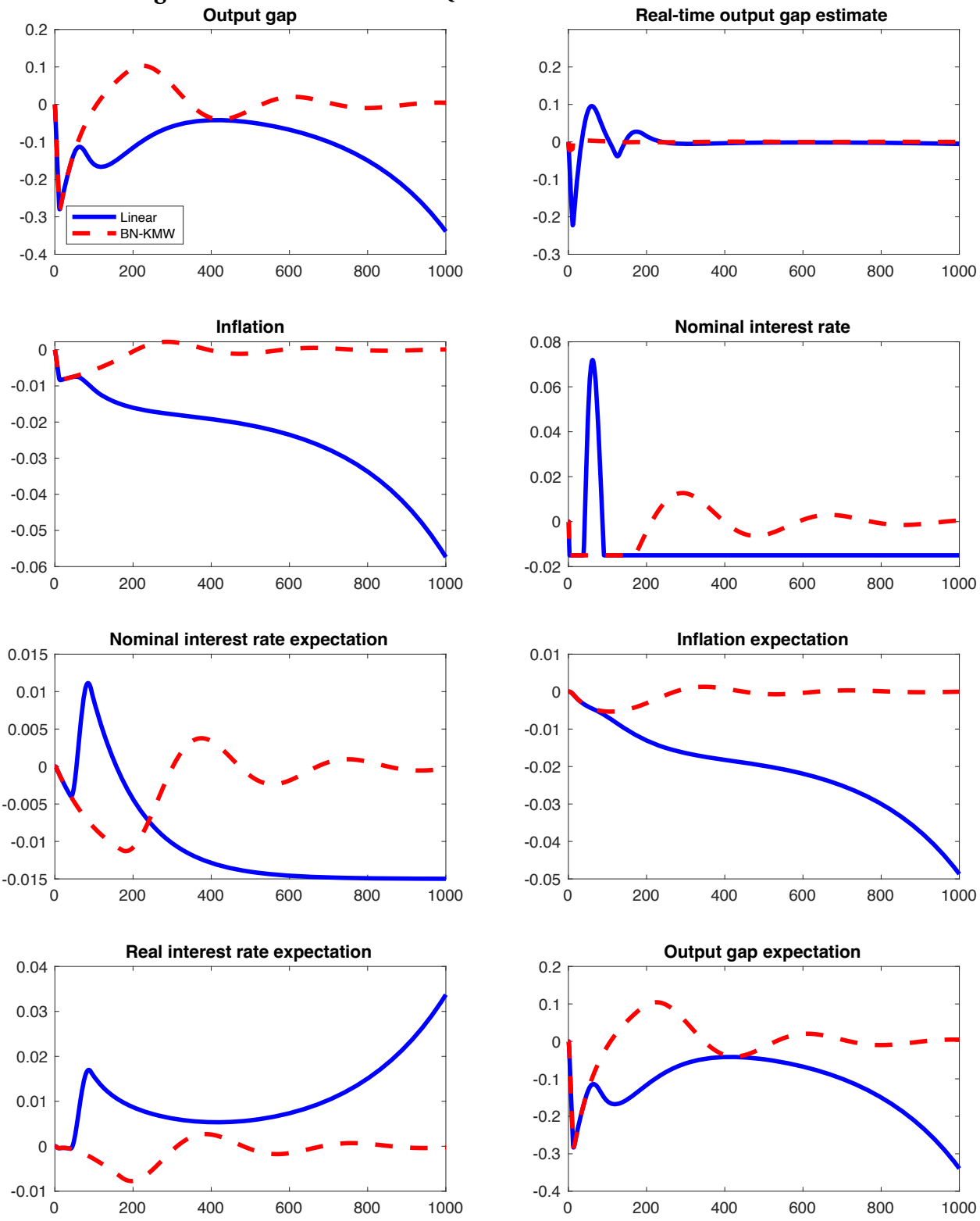
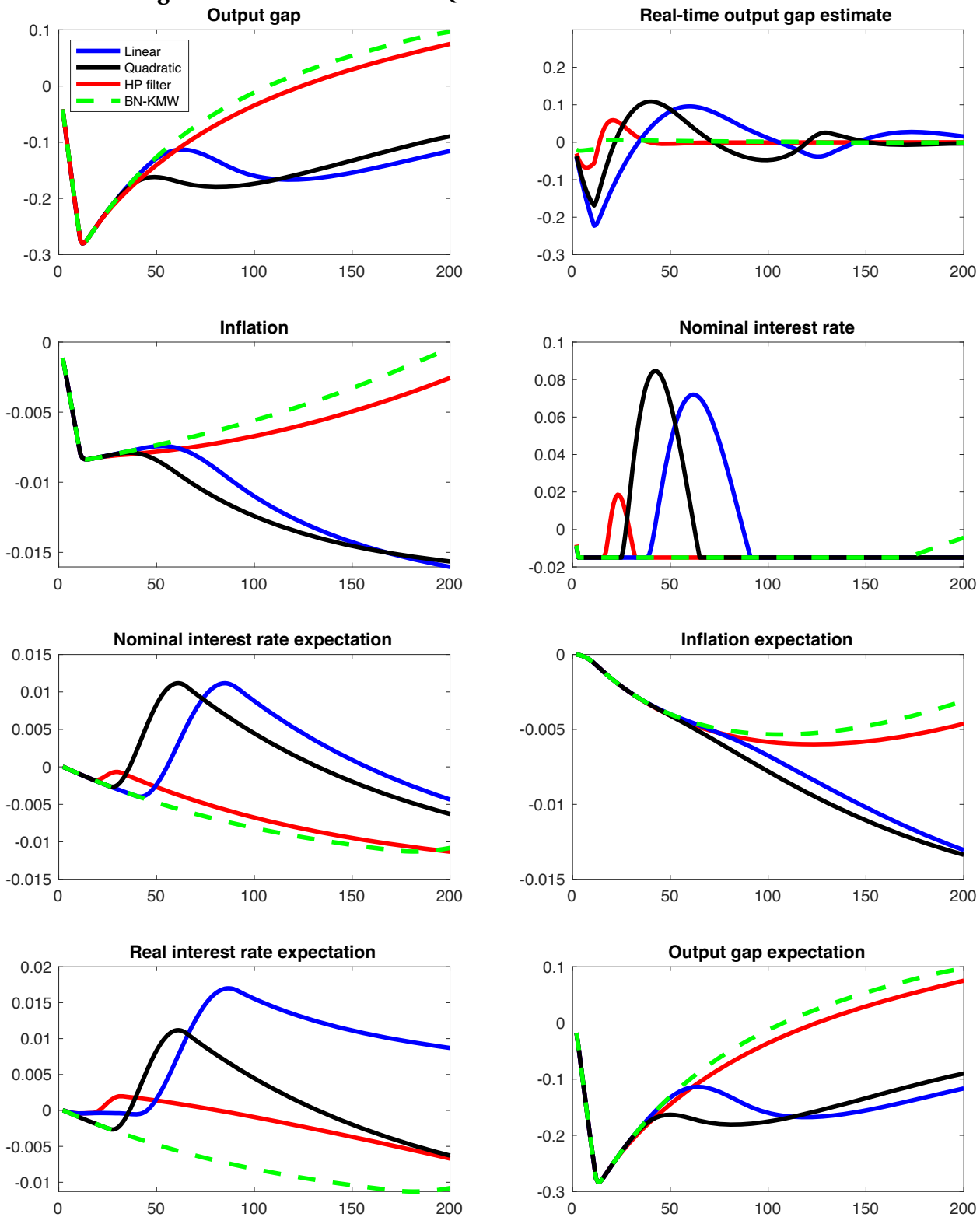


Figure A.8: RESPONSE TO A SEQUENCE OF NEGATIVE DEMAND SHOCKS



G Without ZLB and negative interest rates

This appendix reports the results for two scenarios: no ZLB and -1% lower bound on nominal interest rates.

G.1 Without ZLB

This appendix reports iso-loss contours and statistical properties of revisions to output gap estimates for each detrending method, when there is no lower bound on nominal interest rates.

Figures A.9-A.10 plot iso-loss contours for each detrending method with two different degrees of policy inertia, $\theta_r = 0$ and 0.6, for the case without ZLB. The rank of the size of stability region among detrending methods remains unchanged.

Table A.2 reports optimized policy coefficients and welfare loss for each method. Table A.3 reports 10-year revisions to output gap estimates associated with the optimized monetary policy coefficients with no inertia. The general statistical property is similar to the baseline results. FOD generates the smallest welfare loss; if we extend θ_π to beyond 10, the minimized loss is likely to go down even further. HP, Hamilton, CF band-pass and BN and BN-KMW filters have similar associated minimized welfare losses. Also, the optimal loss is uniformly lower than the case with a ZLB in the baseline for every detrending method.

G.2 Allowing -1% bound on nominal interest rates

After the Great Recession, several central banks around the world, including the Bank of Japan, the Denmarks Nationalbank, the Sveriges Riksbank, the Swiss National Bank and the European Central Bank, have used negative nominal interest rates as monetary policy instruments.[40] This appendix reports the iso-loss contours when $\theta_r = 0$, the size of stability regions and optimal policy when the bound on (annualized) nominal interest rates is -1% (instead of 0%). Tables A.4 and A.5 report the stability results and optimal policy. The results are qualitatively similar to those under a higher inflation target (Appendix J) and the intuition is the same. For every detrending method, the size of stability region increases by allowing negative interest rates. Moreover, a lower volatility of inflation and the OG leads to higher welfare associated with the optimal policy, relative to the baseline results. The improved stability properties also generally allow a larger response to OG estimates under the optimal policy.

[40] Ulate (2021) considers the effectiveness of this policy and shows that monetary policy rate in negative territory is 60% to 90% as effective as in positive territory. Implementing negative nominal interest rates can still facilitate stimulating and stabilizing the economy.

Figure A.9: WITHOUT ZLB: **ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS**

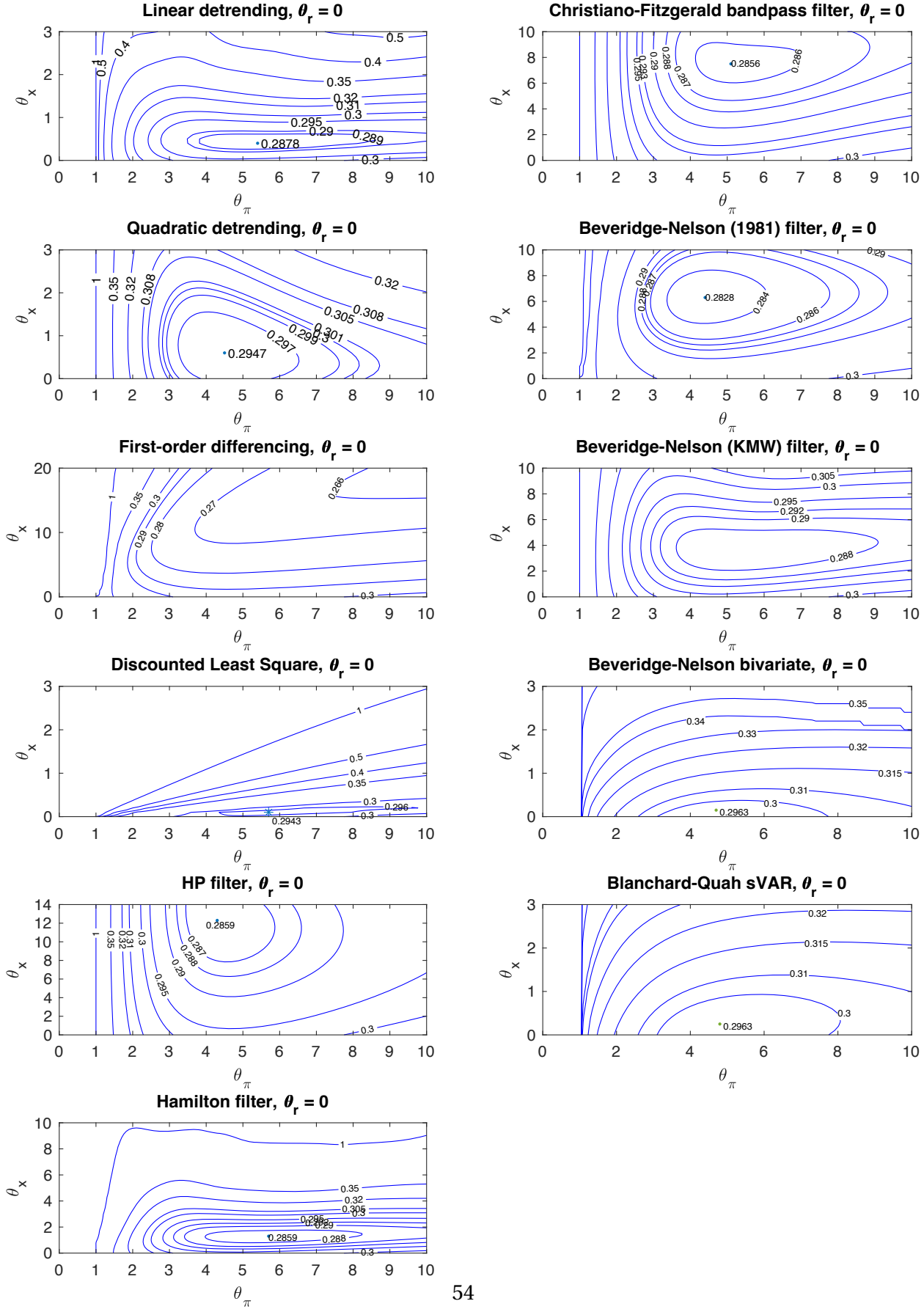


Figure A.10: WITHOUT ZLB: **ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS**

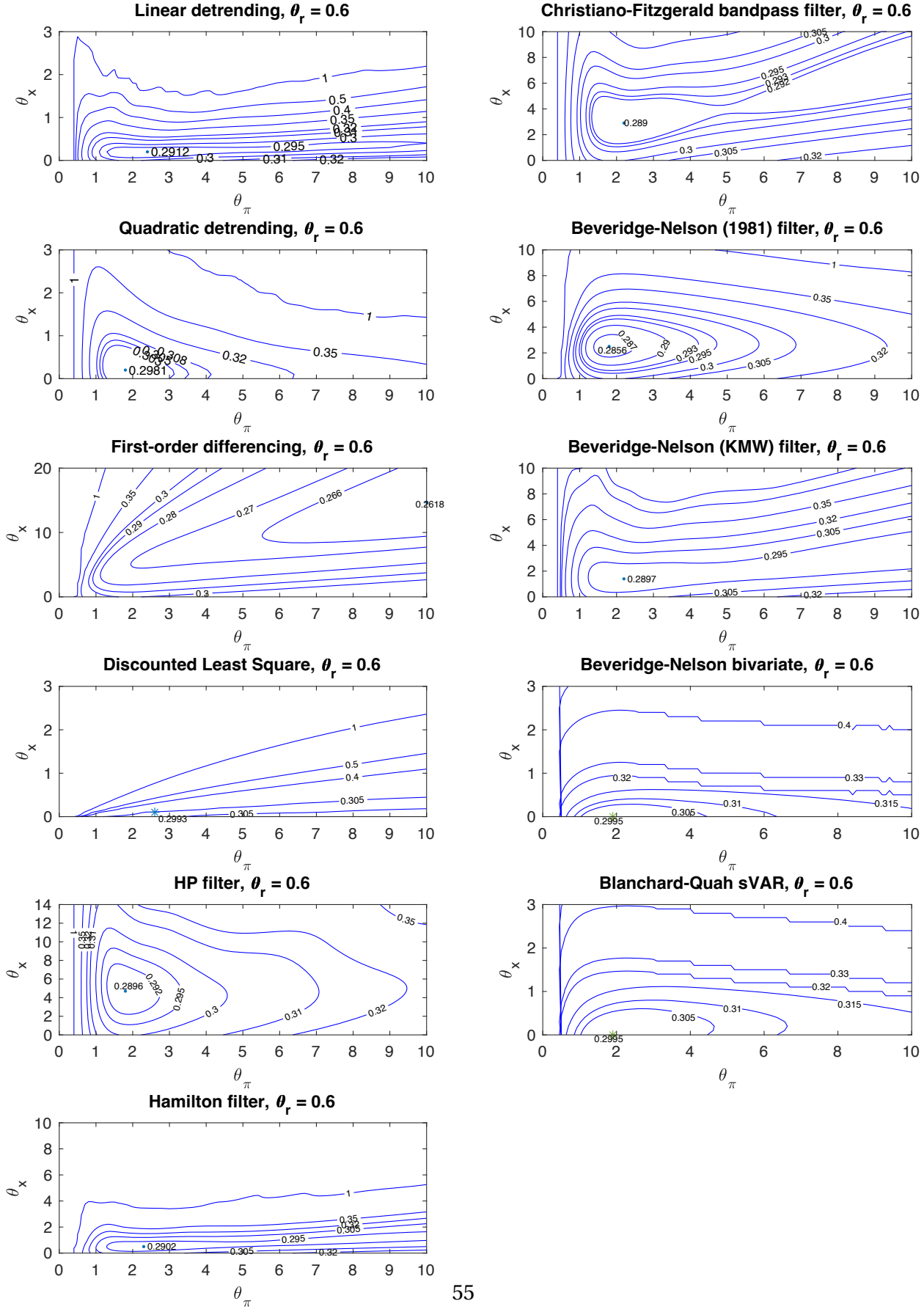


Table A.2: WITHOUT ZERO LOWER BOUND: OPTIMAL POLICY AND WELFARE LOSS

Method	Coeff		SD			Coeff		SD		
	θ_π	θ_x	σ_π	σ_x	Loss	θ_π	θ_x	σ_π	σ_x	Loss
Panel A: methods intrinsically producing smaller $\widetilde{x}_{t-1}$						**Panel B: methods intrinsically producing larger $\widetilde{x}_{t-1}$**				
Christiano-Fitzgerald band-pass filter						Blanchard-Quah				
$\theta_r = 0$	5.1	7.5	0.469	4.674	0.286	$\theta_r = 0$	4.8	0.2	0.472	4.961	0.296
$\theta_r = 0.2$	4.1	6.0	0.469	4.688	0.286	$\theta_r = 0.2$	3.8	0.2	0.472	4.967	0.297
$\theta_r = 0.6$	2.2	2.9	0.470	4.777	0.289	$\theta_r = 0.6$	1.9	0.0	0.473	5.031	0.300
$\theta_r = 0.8$	1.3	1.5	0.471	4.980	0.296	$\theta_r = 0.8$	1.0	0.0	0.475	5.199	0.307
Beveridge-Nelson (KMW)						Beveridge-Nelson (bivariate)					
$\theta_r = 0$	5.3	3.6	0.469	4.706	0.286	$\theta_r = 0$	4.7	0.1	0.472	4.957	0.296
$\theta_r = 0.2$	4.2	2.9	0.469	4.715	0.287	$\theta_r = 0.2$	3.7	0.1	0.472	4.963	0.297
$\theta_r = 0.6$	2.2	1.4	0.470	4.796	0.290	$\theta_r = 0.6$	1.9	0.0	0.473	5.031	0.300
$\theta_r = 0.8$	1.3	0.7	0.471	5.001	0.297	$\theta_r = 0.8$	1.0	0.0	0.475	5.199	0.307
HP filter						Hamilton					
$\theta_r = 0$	4.3	12.3	0.471	4.611	0.286	$\theta_r = 0$	5.7	1.3	0.469	4.738	0.287
$\theta_r = 0.2$	3.5	9.7	0.471	4.631	0.286	$\theta_r = 0.2$	4.5	1.0	0.469	4.748	0.287
$\theta_r = 0.6$	1.8	4.7	0.472	4.717	0.290	$\theta_r = 0.6$	2.3	0.5	0.470	4.821	0.290
$\theta_r = 0.8$	1.0	2.2	0.474	4.938	0.298	$\theta_r = 0.8$	1.4	0.3	0.471	5.035	0.298
First-order differencing						Quadratic trend					
$\theta_r = 0$	10.0	19.4	0.462	4.163	0.265	$\theta_r = 0$	4.5	0.6	0.472	4.894	0.295
$\theta_r = 0.2$	10.0	17.7	0.461	4.176	0.264	$\theta_r = 0.2$	3.6	0.5	0.472	4.905	0.295
$\theta_r = 0.6$	10.0	14.7	0.458	4.173	0.262	$\theta_r = 0.6$	1.8	0.2	0.473	4.970	0.298
$\theta_r = 0.8$	10.0	13.7	0.457	4.262	0.263	$\theta_r = 0.8$	1.0	0.1	0.475	5.166	0.306
Beveridge-Nelson (1981)						Linear trend					
$\theta_r = 0$	4.4	6.3	0.469	4.589	0.283	$\theta_r = 0$	5.4	0.4	0.469	4.763	0.288
$\theta_r = 0.2$	3.5	5.0	0.469	4.597	0.283	$\theta_r = 0.2$	4.3	0.3	0.469	4.776	0.288
$\theta_r = 0.6$	1.8	2.5	0.469	4.662	0.286	$\theta_r = 0.6$	2.4	0.2	0.469	4.864	0.291
$\theta_r = 0.8$	1.0	1.3	0.471	4.834	0.292	$\theta_r = 0.8$	1.4	0.1	0.471	5.055	0.298
						Discounted Least Squares					
						$\theta_r = 0$	5.7	0.1	0.471	4.910	0.294
						$\theta_r = 0.2$	5.0	0.1	0.471	4.929	0.295
						$\theta_r = 0.6$	2.6	0.1	0.473	5.032	0.299
						$\theta_r = 0.8$	1.0	0.0	0.475	5.199	0.307

Notes: This table reports optimized policy coefficients and welfare loss for each detrending method under different degrees of interest rate inertia (θ_r) when the zero lower bound is removed. Panel A shows methods producing smaller OG estimates; Panel B shows methods producing larger OG estimates. The interest rate rule is $i_t = \theta_r i_{t-1} + \theta_\pi \pi_{t-1} + \theta_x \widetilde{x}_{t-1} + e_{i,t}$. SD denotes standard deviation from simulation.

Table A.3: PROPERTIES OF REVISIONS TO OUTPUT GAP ESTIMATE (WITHOUT ZLB)

Method	*Output gap estimate revision*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.649	0.815	3.266	-3.805	0.938
Beveridge-Nelson (KMW, 2016)	0.146	0.184	0.740	-0.691	0.967
HP filter	0.974	1.225	5.008	-5.484	0.970
Beveridge-Nelson (1981)	1.341	1.523	6.841	-8.264	0.790
Blanchard-Quah	1.331	2.236	9.674	-11.572	0.908
Beveridge-Nelson (Bivariate)	1.519	2.459	12.969	-12.545	0.835
Hamilton	0.514	0.662	4.428	-3.922	0.897
Quadratic trend	2.036	2.558	9.093	-9.248	0.997
Linear trend	2.487	3.104	9.870	-11.085	0.999

Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method without the zero lower bound. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

Table A.4: -1% INTEREST RATE BOUND: THE SIZE OF STABILITY REGION FOR DIFFERENT DETRENDING METHODS

Stability region (%)				
	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	83.7	84.6	85.1	84.3
Beveridge-Nelson (KMW, 2016)	78.3	78.7	78.2	77.0
HP filter	75.1	75.4	73.9	71.3
First-order differencing	70.0	71.8	75.4	76.7
Beveridge-Nelson (1981)	68.6	58.6	61.9	62.9
Blanchard-Quah	68.6	60.8	56.1	46.1
Beveridge-Nelson (bivariate)	64.5	55.1	39.6	30.3
Hamilton	39.8	38.7	35.4	32.7
Quadratic trend	32.3	31.5	26.1	19.3
Linear trend	22.9	21.4	16.4	12.1
Discounted Least Squares	9.3	9.4	9.1	8.0

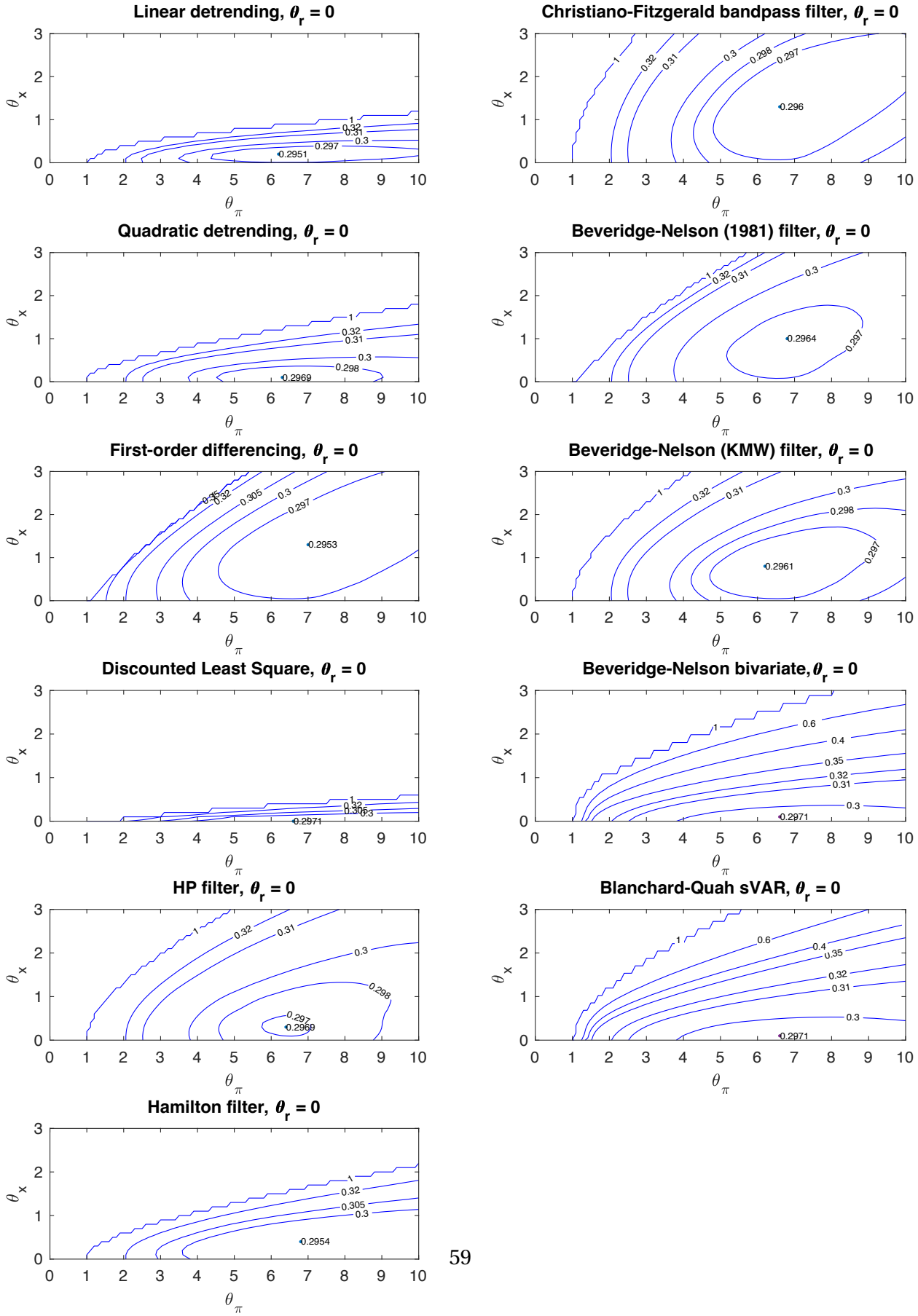
Notes: This table reports the size of the stability region for each detrending method under different degrees of interest rate inertia (θ_r) when the lower bound on nominal interest rates is -1%. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1.

Table A.5: -1% INTEREST RATE BOUND: OPTIMAL POLICY AND WELFARE LOSS

Method	Coeff		SD					Coeff		SD			
	θ_π	θ_x	σ_π	σ_x	Loss	ZLB(%)		θ_π	θ_x	σ_π	σ_x	Loss	ZLB(%)
Panel A: methods intrinsically producing smaller $\widetilde{x}_{t-1}$							**Panel B: methods intrinsically producing larger $\widetilde{x}_{t-1}$**						
Christiano-Fitzgerald band-pass filter							Blanchard-Quah						
$\theta_r = 0$	6.6	1.3	0.475	4.852	0.296	44.3	$\theta_r = 0$	6.6	0.1	0.474	4.902	0.297	42.1
$\theta_r = 0.6$	2.6	0.6	0.474	4.893	0.296	26.8	$\theta_r = 0.6$	2.5	0.0	0.473	4.958	0.298	22.5
Beveridge-Nelson (KMW)							Beveridge-Nelson (bivariate)						
$\theta_r = 0$	6.2	0.8	0.475	4.841	0.296	41.7	$\theta_r = 0$	6.6	0.1	0.474	4.902	0.297	42.1
$\theta_r = 0.6$	2.7	0.5	0.474	4.882	0.296	28.5	$\theta_r = 0.6$	2.5	0.0	0.473	4.958	0.298	22.5
HP filter							Hamilton						
$\theta_r = 0$	6.4	0.3	0.475	4.889	0.297	41.3	$\theta_r = 0$	6.8	0.4	0.475	4.836	0.295	45.7
$\theta_r = 0.6$	2.6	0.2	0.473	4.947	0.298	24.4	$\theta_r = 0.6$	2.8	0.2	0.474	4.881	0.296	29.9
First-order differencing							Quadratic trend						
$\theta_r = 0$	7.0	1.3	0.475	4.811	0.295	50.0	$\theta_r = 0$	6.3	0.1	0.475	4.888	0.297	40.7
$\theta_r = 0.6$	3.1	1.0	0.474	4.796	0.294	38.9	$\theta_r = 0.6$	2.6	0.1	0.474	4.948	0.298	25.2
Beveridge-Nelson (1981)							Linear trend						
$\theta_r = 0$	6.8	1.0	0.475	4.848	0.296	47.7	$\theta_r = 0$	6.2	0.2	0.475	4.809	0.295	41.6
$\theta_r = 0.6$	2.7	0.6	0.474	4.867	0.296	31.0	$\theta_r = 0.6$	2.8	0.1	0.473	4.880	0.296	28.5
							Discounted Least Squares						
							$\theta_r = 0$	6.6	0.0	0.474	4.902	0.297	42.1
							$\theta_r = 0.6$	2.5	0.0	0.473	4.958	0.298	22.5

Notes: This table reports optimized policy coefficients and welfare loss for each detrending method when the lower bound on nominal interest rates is −1%. Panel A shows methods producing smaller OG estimates; Panel B shows methods producing larger OG estimates. SD denotes standard deviation from simulation.

Figure A.11: -1% INTEREST RATE BOUND: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS

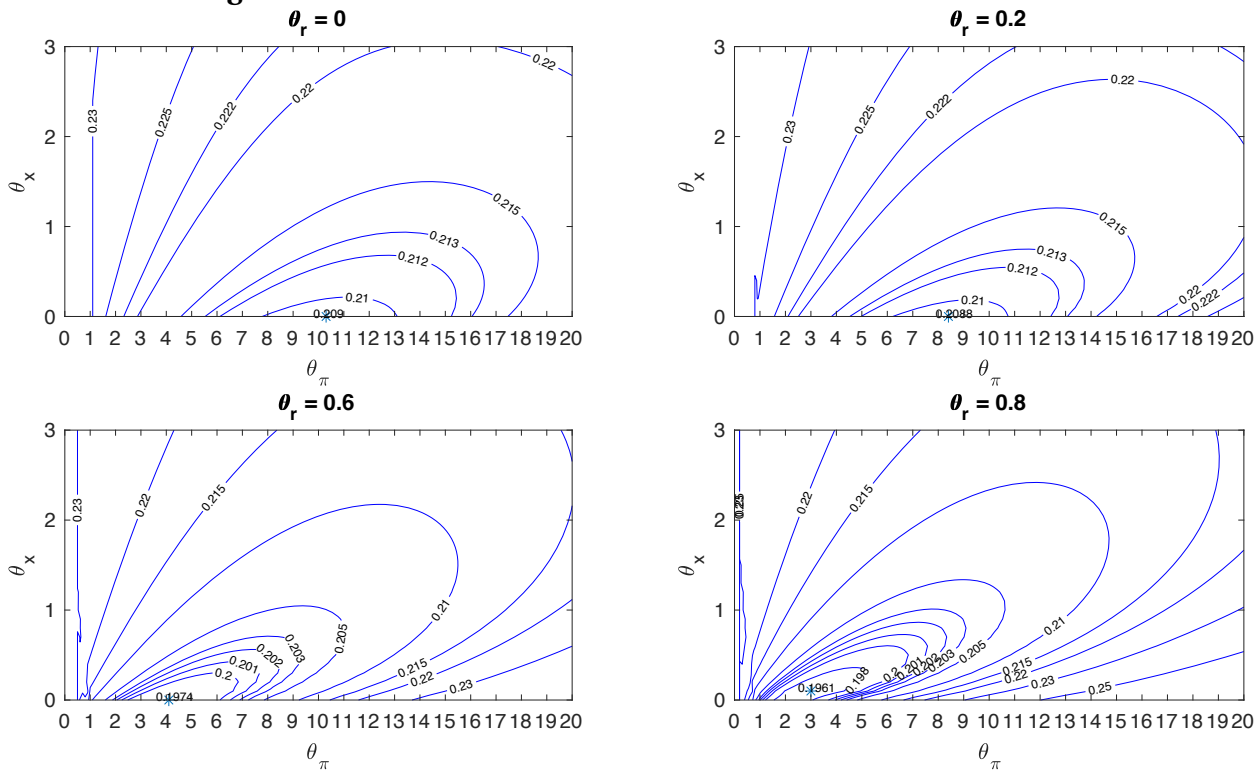


H Optimal policy under full-information Rational Expectation

This appendix shows the iso-loss contours and optimal monetary policy for our baseline model under the assumption that households and the policymaker have full-information Rational Expectation (RE).

As in the learning model, we simulate our model for 1,000 + 40,000 periods for each combination of policy coefficients under the parametrization of the baseline model. The final 40,000 periods are used. Figure A.12 plots the iso-loss contours under RE with different degrees of policy inertia. The minimized welfare loss under RE is smaller than the minimized welfare loss under learning with any of the detrending methods and for all degrees of interest rate inertia (as well as the scenario when the policymaker has knowledge of the true output gap). Having complete knowledge of the model economy, unsurprisingly, improves the policymaker's welfare.

Figure A.12: ISO-LOSS CONTOURS UNDER FULL-INFORMATION RE



I Alternative parameterization and modification of OG estimation methods

This appendix provides some results for alternative parameterization and modification of a few filters. In the baseline case, we used a one-sided HP filter with a smoothing parameter of 1600. Here we consider a smoothing parameter of 650 suggested by Wolf et al. (2026). Figure A.13 shows the iso-loss contours for this alternative parameterization. We find that the size of the stability region for the one-sided HP filter with a smoothing parameter of 650 increases substantially relatively to the HP filter with a smoothing parameter of 1600. The result holds for all values of the interest rate smoothing parameter, i.e., $\theta_r = 0, 0.2, 0.6$, and 0.8.

Additionally, we consider a modification of the Hamilton filter developed in Quast and Wolters (2022). We set the parameters to their suggested values (a lower bound for forecast horizons of 4 and an upper bound of 12). Figure A.14 shows the iso-loss contours for this modification. We find that the modified Hamilton filter has a significantly larger stability region relative to the original Hamilton filter. The result holds for all values of the interest rate smoothing parameter, i.e., $\theta_r = 0, 0.2, 0.6$, and 0.8.

Figure A.13: ISO-LOSS CONTOURS WITH THE HP FILTER (650)

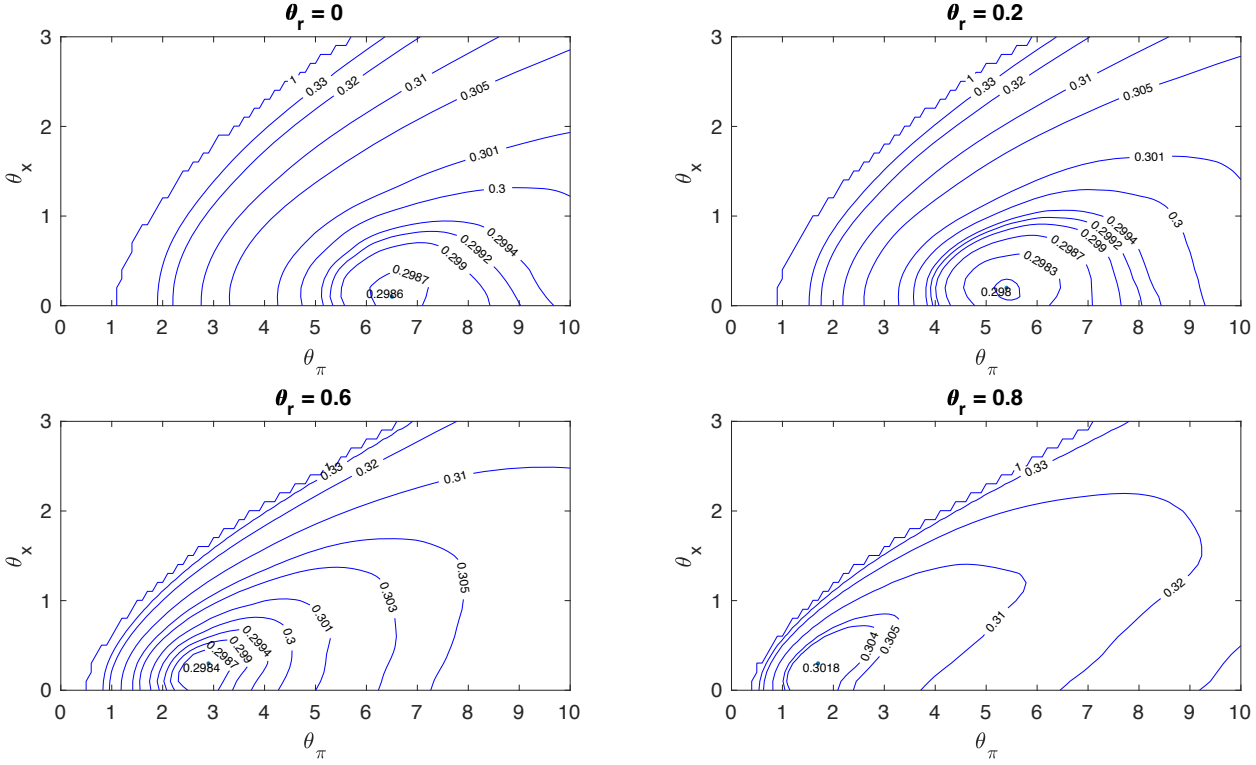
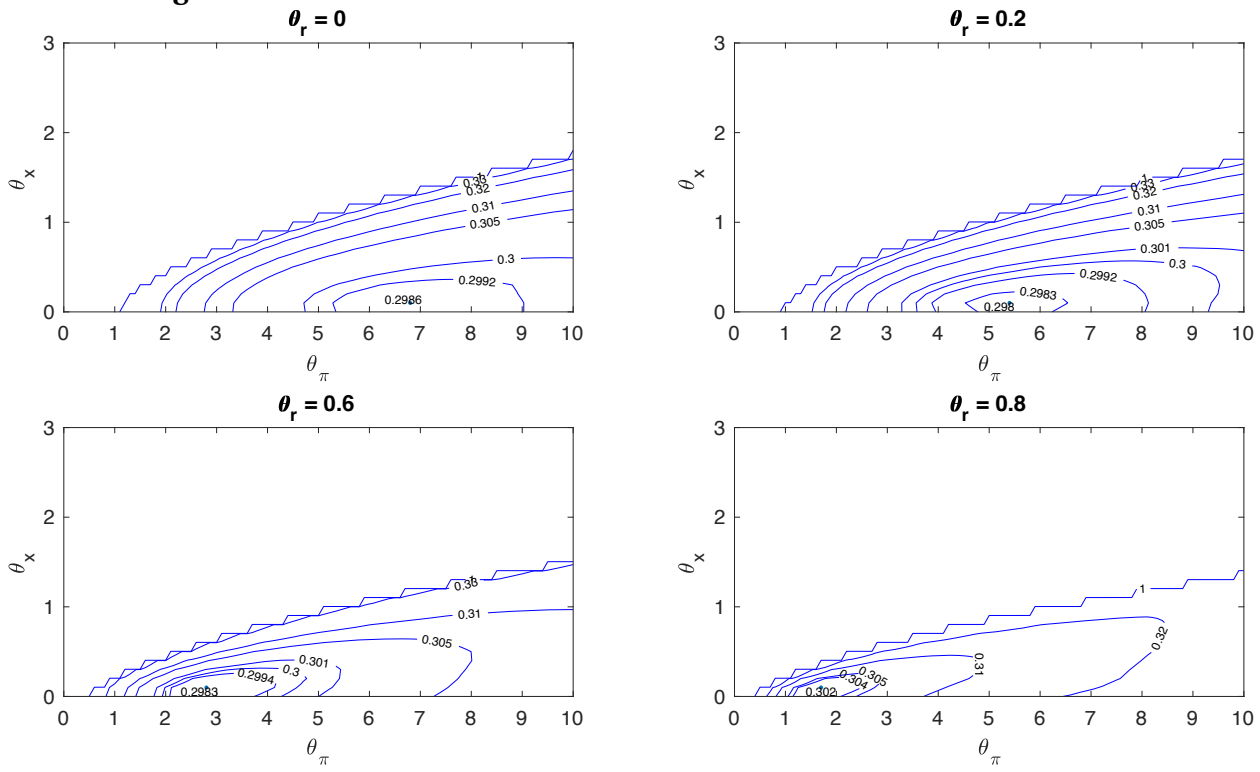


Figure A.14: ISO-LOSS CONTOURS WITH THE MODIFIED HAMILTON FILTER



J A 4% inflation target

There was a consensus among policymakers that inflation should not only stay stable, but also remain low during the period of the Great Moderation (a 2% inflation target was common consensus among central banks). Since the Great Recession, nominal interest rates have remained close to zero which has led to growing concerns over the likelihood of falling into a liquidity trap. In case of adverse shocks, central banks may lack space for expansionary monetary policy. There was popular discussion surrounding inflation targets being raised above 2%: see, for example, Blanchard et al. (2010), Coibion et al. (2012), and L'Huillier and Schoenle (2021). This appendix analyzes a 4% inflation target and compares the results with the baseline case of 2%, reporting stability regions, iso-loss contours and revision properties of output gap estimates.

Table A.6 shows the stability regions and Table A.7 shows the statistics of 10-year output gap estimation revision associated with the optimized monetary policy coefficients with no inertia. Panel A of Table A.8 reports optimal policy coefficients. There is a significant increase in the size of the stability region with every detrending method, comparing with Panel A of Table 1. For instance, without interest rate inertia, the size of the stability region increases from 66.4% to 79.3% for the HP filter, from 79.8% to 85.8% for the CF bandpass filter, from 70.6% to 86.9%

for the BN-KMW filter, and from 53.7% to 67.6% for the BQ method. This suggests a higher inflation target can provide the policymaker more maneuvering space to avoid large welfare losses. The chance of hitting the ZLB decreases and welfare improves with every detrending method, comparing with Table 2 for the 2% case.[41] With a higher inflation target, it is generally optimal to place a higher reaction coefficient on OG estimates.

A higher inflation target leads to a lower chance of hitting the ZLB and improved ability to stabilize the economy. This reduces the volatility of inflation and OG, which in turn reduces the volatility of households' and policymaker's beliefs and hence reduces the likelihood of large deflation and inflation spirals. Thus, a higher inflation target increases the size of the stability region. At the same time, the smaller volatility of inflation and OG reduces the welfare loss associated with optimal policy. This complements the proposed mechanism in the literature that setting a higher inflation target raises the nominal interest rate, providing the policymaker extra room to lower interest rates in face of economic downturns and reducing the probability of falling into a liquidity trap (L'Huillier and Schoenle, 2021; Blanchard et al., 2010).[42]

Table A.6: THE SIZE OF STABILITY REGION FOR DETRENDING METHODS (4% INFLATION TARGET)

Stability region (%)				
	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	85.8	86.8	87.9	87.6
Beveridge-Nelson (KMW, 2016)	86.9	88.6	91.8	92.8
HP filter	79.3	80.0	79.5	77.9
First-order differencing	74.6	76.7	80.2	81.8
Beveridge-Nelson (1981)	74.0	76.0	79.3	80.6
Blanchard-Quah	67.6	61.6	54.7	50.7
Beveridge-Nelson (bivariate)	63.8	50.2	45.4	41.9
Hamilton	50.2	48.7	44.3	41.0
Quadratic trend	40.7	39.7	31.9	23.1
Linear trend	28.9	26.9	20.4	15.1
Discounted Least Squares	3.8	3.8	3.5	3.0

Notes: This table reports the size of the stability region for each detrending method under a 4% inflation target with different degrees of interest rate inertia (θ_r). The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. The ZLB is imposed.

Figure A.15 plots the iso-loss contours for each detrending method for $\theta_r = 0$ when the inflation target is 4%.

[41] The Table only reports values for two degrees of inertia (0 and 0.6) for brevity.

[42] Evans et al. (2022) provide a different rationale for preferring a higher inflation target in the context of a nonlinear NK model.

Figure A.15: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS (4% INFLATION TARGET)

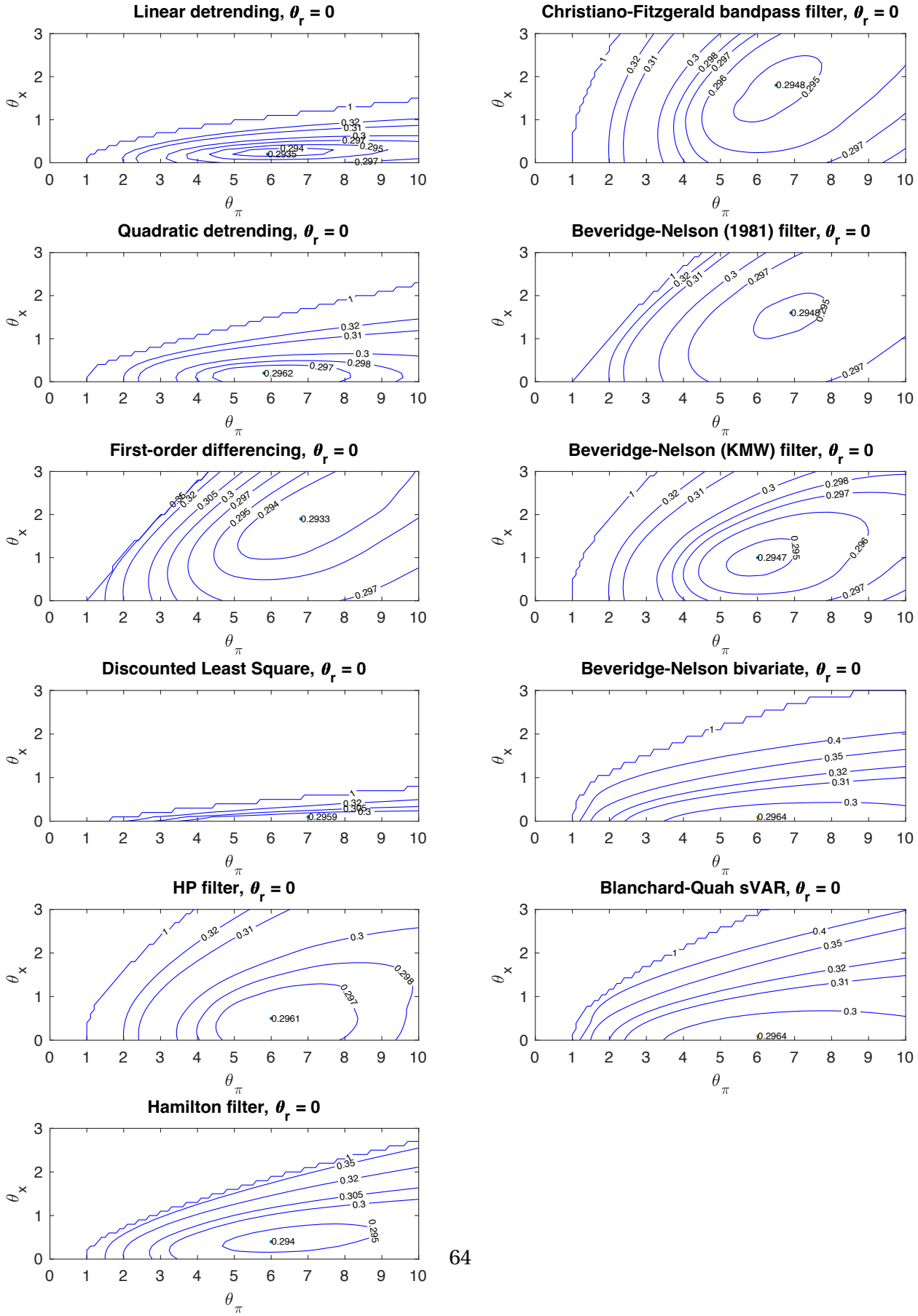


Table A.7: REVISIONS TO OUTPUT GAP ESTIMATE (4% INFLATION TARGET)

Method	*Output gap estimate revisions*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.635	0.799	3.121	-3.493	0.938
Beveridge-Nelson (KMW, 2016)	0.160	0.201	0.697	-0.825	0.973
HP filter	1.015	1.278	5.180	-5.685	0.974
Beveridge-Nelson (1981)	1.331	1.508	6.581	-8.271	0.795
Blanchard-Quah	1.490	2.172	9.638	-11.413	0.906
Beveridge-Nelson (bivariate)	1.503	2.435	12.803	-12.505	0.856
Hamilton	0.551	0.709	4.513	-3.970	0.873
Quadratic trend	2.171	2.726	9.674	-10.011	0.997
Linear trend	2.648	3.305	10.078	-11.822	0.999

Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method under a 4% inflation target. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

K Dual mandate

In a 2012 speech, the then-Chairwoman of the US Federal Reserve Janet Yellen described the Fed's statutory mandate of "maximum employment and price stability" by using a simple welfare loss function that assigns equal weights for annualized inflation and the unemployment gap (difference between actual unemployment and the natural rate of unemployment); see Yellen (2012). Then she transforms the unemployment gap into the OG using Okun's law with a coefficient of 2 (as in our model). Under the dual mandate regime, the welfare loss function is identical to function (2.8) in the main text, but is characterized by a higher relative weight on OG volatility than our baseline parameterization. Following Debortoli et al. (2019), a unit weight on unemployment gap volatility translates into a weight of $\lambda_x^{annual} = 0.25$ on OG variability in an annualized model, which then converts to a quarterly weight of $\lambda_x = 0.25/16 \approx 0.015$ accordingly. This weight on OG variability under the dual mandate is about five times the weight assigned by Woodford (2003) under inflation targeting. A similar approach to the dual mandate with a substantially higher weight on economic activity is also introduced in Svensson (2011).

Note that the only difference between our baseline analysis and the dual mandate is the relative weight on OG variability in the loss function; the remaining aspects of our model are unchanged. As a consequence, the size of the stability region and properties of OG revisions in the dual mandate regime for every detrending method are *identical* to those in our baseline case. This appendix reports the iso-loss contours and statistical properties of revisions to output gap estimates for each detrending method under the dual mandate regime.

Table A.8: OPTIMIZED POLICY COEFFICIENTS AND WELFARE LOSS: 4% INFLATION TARGET OR DUAL MANDATE

Panel A: 4% inflation target							Panel B: Dual mandate						
Method	Coeff		SD					Coeff		SD			
	θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)		θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)
Part 1: Methods intrinsically producing smaller and less volatile $\widetilde{x}_{t-1}$													
Christiano-Fitzgerald band-pass filter													
$\theta_r = 0$	6.5	1.8	0.474	4.847	0.295	40.0	$\theta_r = 0$	5.0	1.3	0.481	4.800	0.591	43.0
$\theta_r = 0.6$	2.6	0.7	0.473	4.902	0.295	22.8	$\theta_r = 0.6$	2.2	0.7	0.479	4.839	0.595	31.1
Beveridge-Nelson (Kamber, Morley and Wong, 2018)													
$\theta_r = 0$	6.0	1.0	0.474	4.833	0.295	35.8	$\theta_r = 0$	4.7	0.8	0.482	4.789	0.590	40.1
$\theta_r = 0.6$	2.6	0.5	0.473	4.893	0.295	22.5	$\theta_r = 0.6$	2.1	0.5	0.480	4.818	0.593	29.4
HP filter													
$\theta_r = 0$	6.0	0.5	0.474	4.892	0.296	34.0	$\theta_r = 0$	4.5	0.4	0.481	4.826	0.595	36.2
$\theta_r = 0.6$	2.4	0.3	0.473	4.947	0.297	18.7	$\theta_r = 0.6$	2.2	0.4	0.478	4.871	0.600	29.8
First-order differencing													
$\theta_r = 0$	6.8	1.9	0.474	4.774	0.293	46.8	$\theta_r = 0$	5.2	1.0	0.482	4.772	0.588	48.1
$\theta_r = 0.6$	2.9	1.2	0.473	4.761	0.292	34.2	$\theta_r = 0.6$	2.5	0.8	0.481	4.764	0.586	40.8
Beveridge-Nelson (1981)													
$\theta_r = 0$	6.9	1.6	0.474	4.835	0.295	45.2	$\theta_r = 0$	4.8	0.7	0.482	4.800	0.592	43.3
$\theta_r = 0.6$	2.6	0.8	0.473	4.850	0.294	27.2	$\theta_r = 0.6$	2.3	0.6	0.480	4.815	0.593	35.9
Part 2: Methods intrinsically producing larger and more volatile $\widetilde{x}_{t-1}$													
Blanchard-Quah													
$\theta_r = 0$	6.0	0.1	0.473	4.909	0.296	33.0	$\theta_r = 0$	4.7	0.1	0.479	4.841	0.596	36.8
$\theta_r = 0.6$	2.4	0.0	0.473	4.977	0.298	16.7	$\theta_r = 0.6$	1.9	0.0	0.478	4.897	0.603	20.3
Beveridge-Nelson (bivariate)													
$\theta_r = 0$	6.0	0.1	0.473	4.909	0.296	33.0	$\theta_r = 0$	4.7	0.1	0.479	4.841	0.596	36.8
$\theta_r = 0.6$	2.4	0.0	0.473	4.977	0.298	16.7	$\theta_r = 0.6$	1.9	0.0	0.478	4.897	0.603	20.3
Hamilton													
$\theta_r = 0$	6.0	0.4	0.474	4.819	0.294	35.8	$\theta_r = 0$	5.0	0.4	0.482	4.778	0.589	43.4
$\theta_r = 0.6$	2.6	0.2	0.473	4.880	0.295	23.1	$\theta_r = 0.6$	2.2	0.2	0.479	4.823	0.593	31.0
Quadratic trend													
$\theta_r = 0$	5.8	0.2	0.474	4.889	0.296	32.7	$\theta_r = 0$	4.5	0.1	0.480	4.828	0.595	35.6
$\theta_r = 0.6$	2.5	0.1	0.473	4.963	0.297	19.4	$\theta_r = 0.6$	2.1	0.1	0.478	4.892	0.602	25.8
Linear trend													
$\theta_r = 0$	5.9	0.2	0.474	4.807	0.294	34.1	$\theta_r = 0$	4.6	0.2	0.482	4.762	0.587	39.0
$\theta_r = 0.6$	2.6	0.1	0.472	4.872	0.294	21.5	$\theta_r = 0.6$	2.1	0.1	0.480	4.823	0.594	27.8
Discounted Least Squares													
$\theta_r = 0$	7.0	0.1	0.474	4.872	0.296	40.6	$\theta_r = 0$	5.4	0.1	0.483	4.805	0.594	46.0
$\theta_r = 0.6$	2.4	0.0	0.473	4.977	0.298	16.7	$\theta_r = 0.6$	1.9	0.0	0.478	4.897	0.603	20.3

Notes: This table reports optimized policy coefficients and welfare loss for each detrending method. Panel A presents results under a 4% inflation target; Panel B presents results under a dual mandate. SD denotes standard deviation from simulation.

Figure A.16 plots the iso-loss contours for each detrending method when $\theta_r = 0$, under the dual mandate regime.

Panel B of Table A.8 reports optimized policy coefficients and welfare loss under this scenario with two different degrees of policy inertia, $\theta_r = 0$ and 0.6. Table A.9 reports 10-year output gap revision statistics for each method associated with the optimized policy coefficients with no inertia. Generally, with a higher relative weight on OG volatility, the optimal policy requires an increase in the OG coefficient (θ_x) and a reduction in the inflation coefficient (θ_π). Consequently, inflation volatility increases and OG volatility declines. Note that the decline in optimized θ_π is larger than the increase in θ_x - this contributes to less volatile interest rates in the model and lower chance of hitting the ZLB.

Table A.9: DUAL MANDATE: REVISIONS TO OUTPUT ESTIMATE (2% INFLATION TARGET)

Method	*Output gap estimate revisions*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.633	0.797	3.103	-3.476	0.938
Beveridge-Nelson (KMW, 2016)	0.159	0.199	0.686	-0.826	0.973
HP filter	1.014	1.276	5.174	-5.677	0.974
Beveridge-Nelson (1981)	1.329	1.504	6.538	-8.251	0.803
Blanchard-Quah	1.470	2.151	9.537	-11.154	0.904
Beveridge-Nelson (bivariate)	1.476	2.397	12.816	-12.316	0.880
Hamilton	0.548	0.704	4.469	-3.953	0.872
Quadratic trend	2.176	2.733	9.722	-10.119	0.997
Linear trend	2.612	3.261	9.871	-11.539	0.999

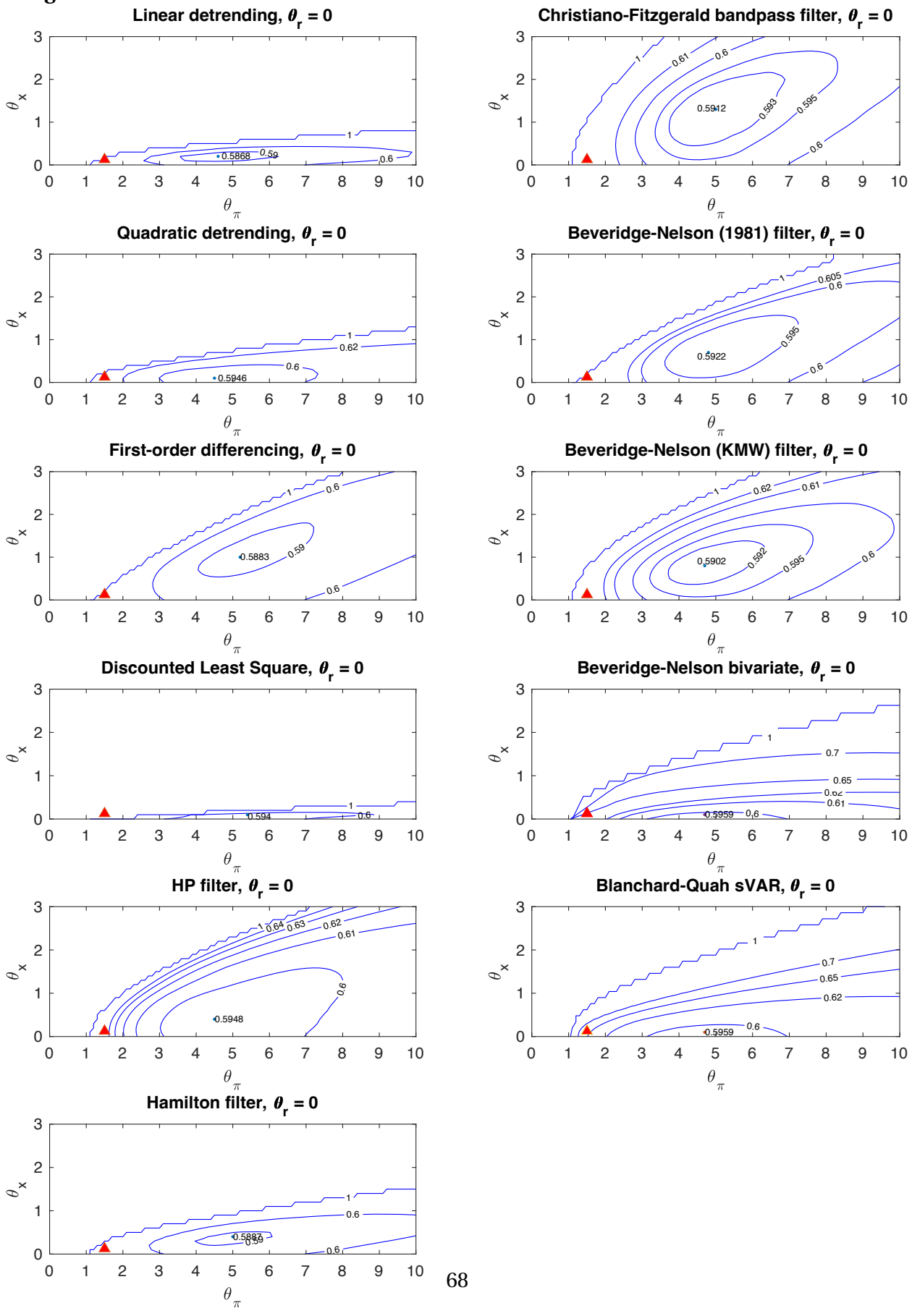
Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method under a dual mandate with a 2% inflation target. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

L Average inflation targeting

In August 2020, the Chairman of US Fed Jeremy Powell publicly announced a revision of the Fed's long-term monetary policy goal to achieve inflation that averages 2% over time, see Powell (2020). There has been interest around this policy and we thus consider the following average inflation targeting (AIT) policy scheme:

$$i_t = \max\{\theta_\pi \bar{\pi}_{t-1}^{AIT} + \theta_x \widetilde{x}_{t-1} + e_{i,t}, -(\bar{\pi}^* + r^*)\}, \tag{A.1}$$

Figure A.16: DUAL MANDATE: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS



as in e.g., Budianto et al. (2023). $\bar{\pi}^*$ is the long-run inflation target and $\bar{\pi}_{t-1}^{AIT}$ is the exponential average of inflation,

$$\bar{\pi}_t^{AIT} = (1-w)\bar{\pi}_{t-1}^{AIT} + w\pi_t, \tag{A.2}$$

and $w \in [0,1]$ is the weight on the new inflation observation and $(1-w)$ is the weight on historical (exponential) average inflation. If $w = 1$, monetary policy reverts to the baseline period-by-period inflation targeting. For illustrative purposes, we set w to 0.2 and the long-run inflation target to 2%. The results under AIT are compared with the case of inflation targeting when there is no inertia in interest rate setting. This appendix illustrates iso-loss contours and other properties for different detrending methods under the AIT rule. Two cases are considered here: the baseline stochastic underlying trend and a deterministic underlying trend.

L.1 Stochastic trend (baseline)

Table A.10 reports the size of stability regions, optimized coefficients, standard deviations, welfare loss and the chance of hitting the ZLB under AIT. Table A.11 reports the properties of 10-year revisions to output gap estimates under AIT with a stochastic trend and the optimized policy coefficients. The general statistical property is similar to the baseline results. Under AIT, the interest rate set by the policymaker no longer responds to period-by-period inflation but to the average inflation (with a high weight on historical inflation average) over a long period of time. The interest rates fluctuate more around the steady state and are less likely to hit the ZLB, which reduces the volatility in interest rates. The ZLB problem is alleviated, which reduces the volatility of agents' beliefs and the likelihood of deflation spirals. This dampens the mechanism that generates self-reinforcing deflation spirals. Thus, the AIT rule increases the size of the stability region with every detrending method. Moreover, the alleviation of the ZLB problem leads to an increase in welfare and an increase in the optimal response to OG estimates.

When θ_x is small, it can be shown that the AIT rule is very close to the baseline Taylor rule with interest rate inertia. Particularly, if $\theta_x = 0$, $1-w$ corresponds exactly to θ_r.[43]

Figure A.17 plots iso-loss contours for each detrending method under AIT. The rank of the size of stability regions among detrending methods is unchanged.

[43] A detailed analysis of the AIT rule and related debates is beyond the scope of the paper and we leave this for future work.

Table A.10: AVERAGE INFLATION TARGETING: STABILITY, OPTIMIZED POLICY COEFFICIENTS, WELFARE LOSS

Average inflation targeting							
		Coeff		SD			
Method	Stability(%)	θ_π	θ_x	σ_π	σ_x	Loss	ZLB(%)
Christiano-Fitzgerald	82.4	5.5	1.3	0.473	4.926	0.297	25.1
Beveridge-Nelson (KMW)	76.1	5.6	1.0	0.474	4.906	0.296	26.3
HP filter	72.5	5.1	0.4	0.473	4.968	0.298	18.5
First-order differencing	68.4	7.7	2.5	0.474	4.806	0.294	50.9
Beveridge-Nelson (1981)	67.0	6.1	1.4	0.474	4.882	0.296	36.3
Blanchard-Quah	63.8	5.2	0.1	0.473	4.989	0.299	17.9
Beveridge-Nelson (bivariate)	59.4	5.2	0.1	0.473	4.989	0.299	17.9
Hamilton	37.0	5.4	0.3	0.473	4.920	0.297	22.5
Quadratic trend	30.2	5.2	0.2	0.473	4.961	0.298	20.1
Linear trend	20.3	5.7	0.2	0.473	4.895	0.296	24.0
Discounted Least Squares	9.0	6.3	0.1	0.474	4.953	0.298	26.5

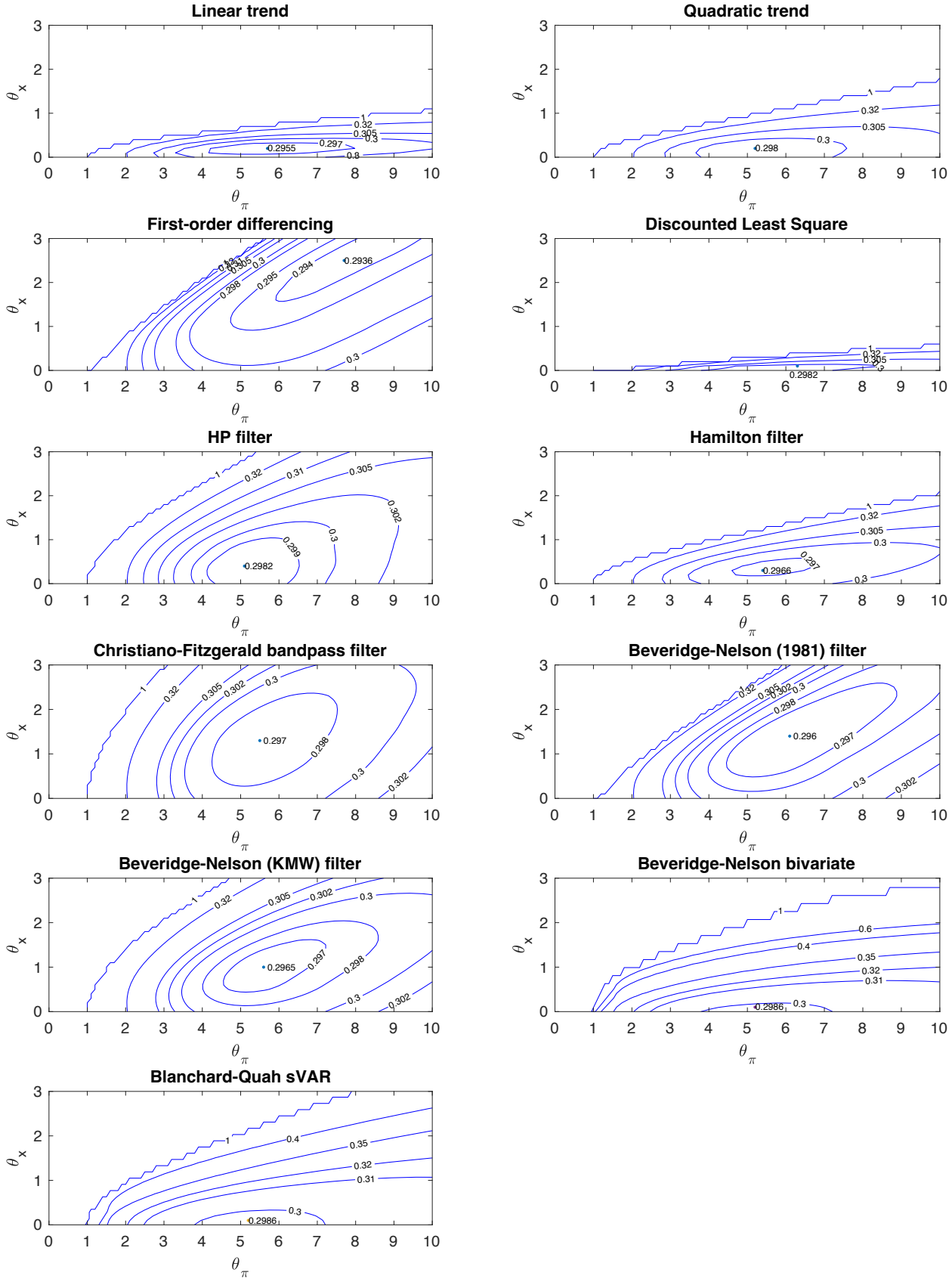
Notes: This table reports the size of the stability region, optimized policy coefficients, and welfare loss for each detrending method under average inflation targeting (AIT) with a stochastic trend. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. The ZLB is imposed.

Table A.11: AVERAGE INFLATION TARGETING: REVISIONS TO OUTPUT GAP ESTIMATE

Method	*Output gap estimate revisions*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.634	0.798	3.118	-3.478	0.938
Beveridge-Nelson (KMW, 2016)	0.160	0.201	0.704	-0.828	0.973
HP filter	1.014	1.277	5.163	-5.680	0.974
Beveridge-Nelson (1981)	1.331	1.507	6.540	-8.284	0.801
Blanchard-Quah	1.510	1.907	9.677	-11.319	0.907
Beveridge-Nelson (bivariate)	1.510	1.946	12.844	-12.467	0.834
Hamilton	0.555	0.715	4.527	-4.010	0.874
Quadratic trend	2.166	2.720	9.681	-10.055	0.997
Linear trend	2.660	3.323	10.144	-11.940	0.999

Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method under average inflation targeting with a stochastic trend. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

Figure A.17: AVERAGE INFLATION TARGETING: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS



L.2 Deterministic trend

Table A.12 reports the size of stability regions, optimized coefficients, standard deviations, welfare loss and the chance of hitting ZLB under AIT with a deterministic underlying trend (see also Appendix M.1 below). Compared with the baseline results, AIT significantly reduces the probability of hitting ZLB and lowers the welfare loss across all detrending methods. Table A.13 reports 10-year revisions to output gap estimates under AIT with a deterministic trend and the optimized policy coefficients. The general statistical property is similar to the baseline results.

Table A.12: AVERAGE INFLATION TARGETING: OPTIMIZED POLICY COEFFICIENTS, WELFARE LOSS AND STABILITY, WITH A DETERMINISTIC TREND

Average inflation targeting							
		Coeff		SD			
Method	*Stability(%)*	θ_π	θ_x	σ_π	σ_x	*Loss*	*ZLB(%)*
Christiano-Fitzgerald	82.6	4.9	2.1	0.475	4.788	0.295	29.286
Beveridge-Nelson (KMW)	78.1	4.8	1.4	0.475	4.772	0.294	27.153
HP filter	72.8	4.1	0.5	0.475	4.859	0.297	14.136
Beveridge-Nelson (1981)	63.7	4.8	0.9	0.475	4.827	0.296	26.282
Blanchard-Quah	59.5	4.2	0.2	0.475	4.884	0.297	13.303
Beveridge-Nelson (bivariate)	55.4	4.2	0.2	0.475	4.883	0.297	13.283
Hamilton	41.6	4.7	0.5	0.475	4.772	0.294	25.454
Quadratic trend	38.1	4.0	0.2	0.475	4.841	0.296	13.381
Linear trend	28.0	4.6	0.3	0.475	4.726	0.292	22.459
Discounted Least Squares	34.2	4.1	0.0	0.475	4.886	0.297	11.748

Notes: This table reports the size of the stability region, optimized policy coefficients, and welfare loss for each detrending method under average inflation targeting with a deterministic trend. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1.

Figure A.18 plots the iso-loss contours for each detrending method under AIT with a deterministic trend. As can be seen from the figures, the rank of the stability regions among detrending methods continues to remain robust.

M Alternative assumptions about the trend

Policymakers and economists are often uncertain if the true underlying trend process is deterministic or stochastic. Identifying the "true" nature of output trend, i.e. stochastic v.s. deterministic, has always been an open question (Nelson and Plosser, 1982; Gordon, 1997). Moreover, policymakers in real time may be uncertain about a potential acceleration or slowdown in long-run growth rates. This is one frequently adopted way to capture the acceleration of productivity

Figure A.18: AVERAGE INFLATION TARGETING: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS, WITH A DETERMINISTIC TREND

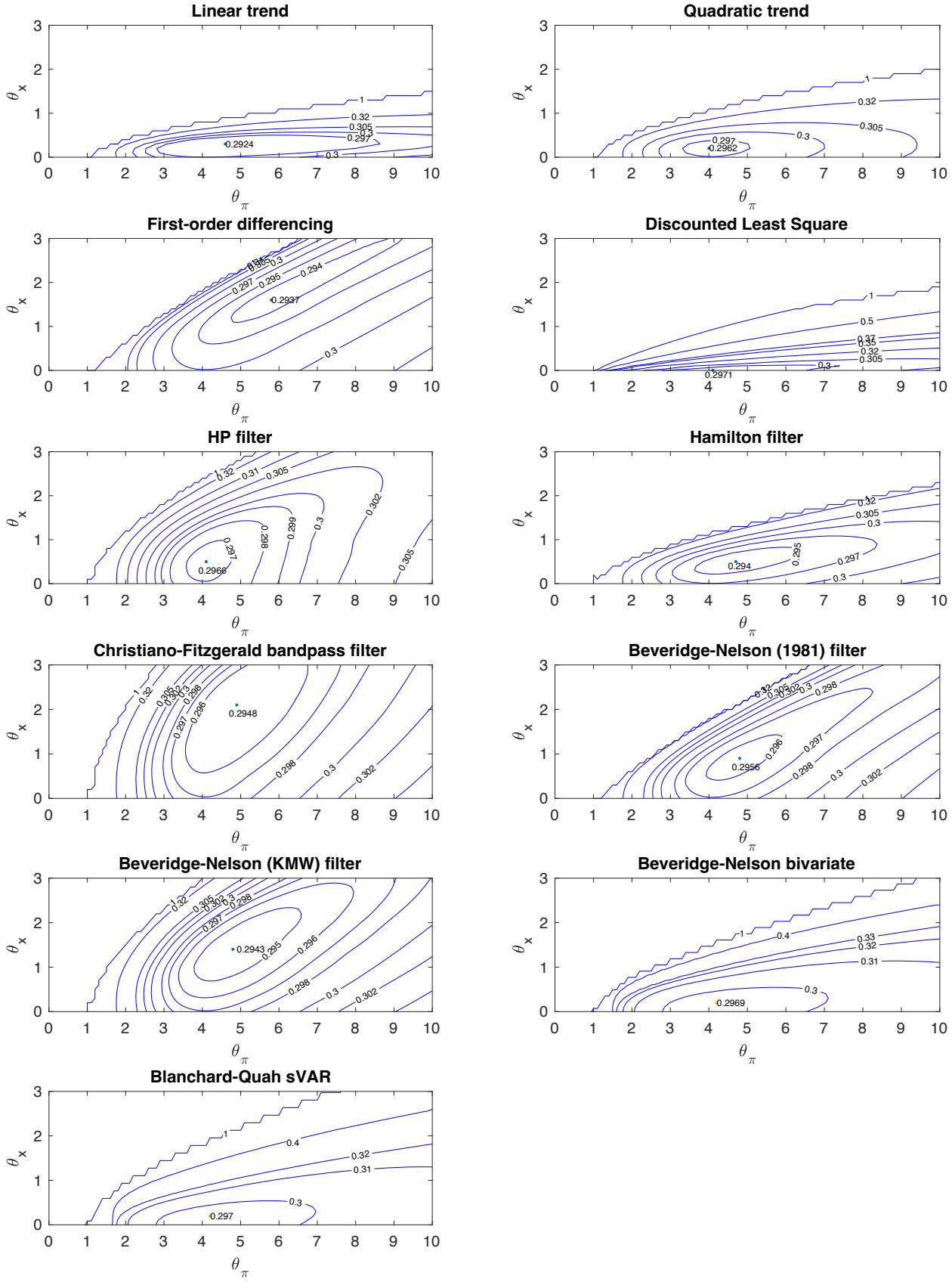


Table A.13: AVERAGE INFLATION TARGETING: REVISIONS TO OUTPUT GAP ESTIMATE, WITH A DETERMINISTIC TREND

Method	*Output gap estimate revisions*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.567	0.712	2.919	-3.161	0.933
Beveridge-Nelson (KMW, 2016)	0.114	0.144	0.671	-0.638	0.944
HP filter	0.896	1.127	4.554	-4.930	0.971
Beveridge-Nelson (1981)	1.323	1.490	6.170	-6.738	0.845
Blanchard-Quah	1.699	2.659	9.600	-11.375	0.902
Beveridge-Nelson (bivariate)	1.512	2.534	10.534	-15.240	0.880
Hamilton	0.458	0.589	3.226	-4.433	0.853
Quadratic trend	1.837	2.311	8.666	-9.339	0.997
Linear trend	2.117	2.658	8.334	-10.903	0.999

Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method under average inflation targeting with a deterministic trend. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

growth during the IT boom in the late 1990s or the slowdown in the decade after the 2007–08 Global Financial Crisis (e.g., Holston et al., 2017). This appendix provides full numerical results under two alternative assumptions (the underlying trend is deterministic, or contains shocks to both its level and growth rate) and characterizes robust monetary policy under uncertainty about the trend. Our main results are robust to both alternative assumptions; in particular, the ranking of detrending methods in terms of the stability region is unchanged.

M.1 A deterministic underlying trend

In lieu of the stochastic trend productivity process (2.3)-(2.4), this section considers an underlying deterministic output trend as follows

$$a_t^p = a_0^p + g\,t + e_{d,t}, \tag{A.1}$$

$$\Delta y_t^p = \Delta a_t^p, \tag{A.2}$$

and re-evaluates the performance of detrending methods and interest rate decisions. a_0^p is the initial value of productivity, g is trend growth rate, t denotes the passage of time (quarter) and $e_{d,t} \sim IIN(0,\sigma_d^2)$ is the innovation to the trend productivity. σ_d is calibrated to the same value as σ_a in the baseline.

Panel A of Table A.14 reports the size of the stability region associated with each detrending method, under different degrees of inertia when the underlying trend is deterministic. Table A.15 reports 10-year revisions to output gap estimates with a deterministic trend and the optimized

policy coefficients with no inertia. The majority of detrending methods (except FOD, HP and BN (1981) filters) are associated with a larger stability region, compared with Table 1. But the ranking of detrending methods in terms of the stability region is unchanged. Panel A of Table A.16 demonstrates the optimized monetary policy coefficients, associated welfare loss and the proportion of simulation periods hitting the ZLB. Comparing with the case of a stochastic trend (Table 2), the optimized coefficient on output gap estimates θ_x becomes larger. For instance, well-performing methods, such as FOD, CF band-pass, and BN-KMW filter continue to have smaller welfare loss. They produce smaller output gap estimates (similar to the illustration in Figure A.4) and require a larger θ_x to achieve the minimized welfare loss than the remaining methods.

Table A.14: THE SIZE OF STABILITY REGION UNDER ALTERNATIVE TREND ASSUMPTIONS

Stability region (%)	Panel A: Deterministic trend				Panel B: Shocks to level and growth			
	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	79.9	81.0	80.9	79.1	78.3	79.0	78.5	76.9
Beveridge-Nelson (KMW, 2016)	73.3	74.2	72.9	68.7	68.5	68.5	67.3	65.5
HP filter	66.2	66.0	62.1	56.3	64.3	64.0	60.6	56.4
First-order differencing	51.7	54.6	58.8	51.3	57.7	60.3	63.6	64.5
Beveridge-Nelson (1981)	52.7	55.6	59.8	59.4	55.1	57.4	60.8	61.9
Blanchard-Quah	51.8	47.6	41.2	35.3	50.9	43.2	35.9	32.0
Beveridge-Nelson (bivariate)	44.2	37.9	33.0	30.0	42.6	33.7	32.3	24.1
Hamilton	33.3	32.3	28.4	24.8	27.0	26.6	24.6	22.8
Quadratic trend	31.1	29.3	22.2	15.8	6.5	6.4	5.8	4.5
Linear trend	21.8	20.5	16.1	12.4	4.5	4.3	3.5	2.7
Discounted Least Squares	1.7	1.8	1.8	1.7	1.7	1.7	1.6	1.6

Notes: This table reports the size of the stability region for each detrending method under alternative trend assumptions. Panel A assumes a deterministic trend; Panel B assumes shocks to both the level and growth rate of the trend. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. The ZLB is imposed.

Figure A.19 plots the iso-loss contours for each detrending method for $\theta_r = 0$ when the true underlying trend is deterministic.

M.2 Both level and growth shocks to the trend

While the baseline results are obtained with only shocks to the level of the trend, this section analyzes the case when there are both level and growth shocks to the trend, as in e.g., Harvey (1985), Stock and Watson (1998). This captures shifts in the long-run growth rates in the data, such as the growth slowdown in the 1970s and acceleration in the 1990s. The trend productivity

Figure A.19: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS, WITH A DETERMINISTIC TREND

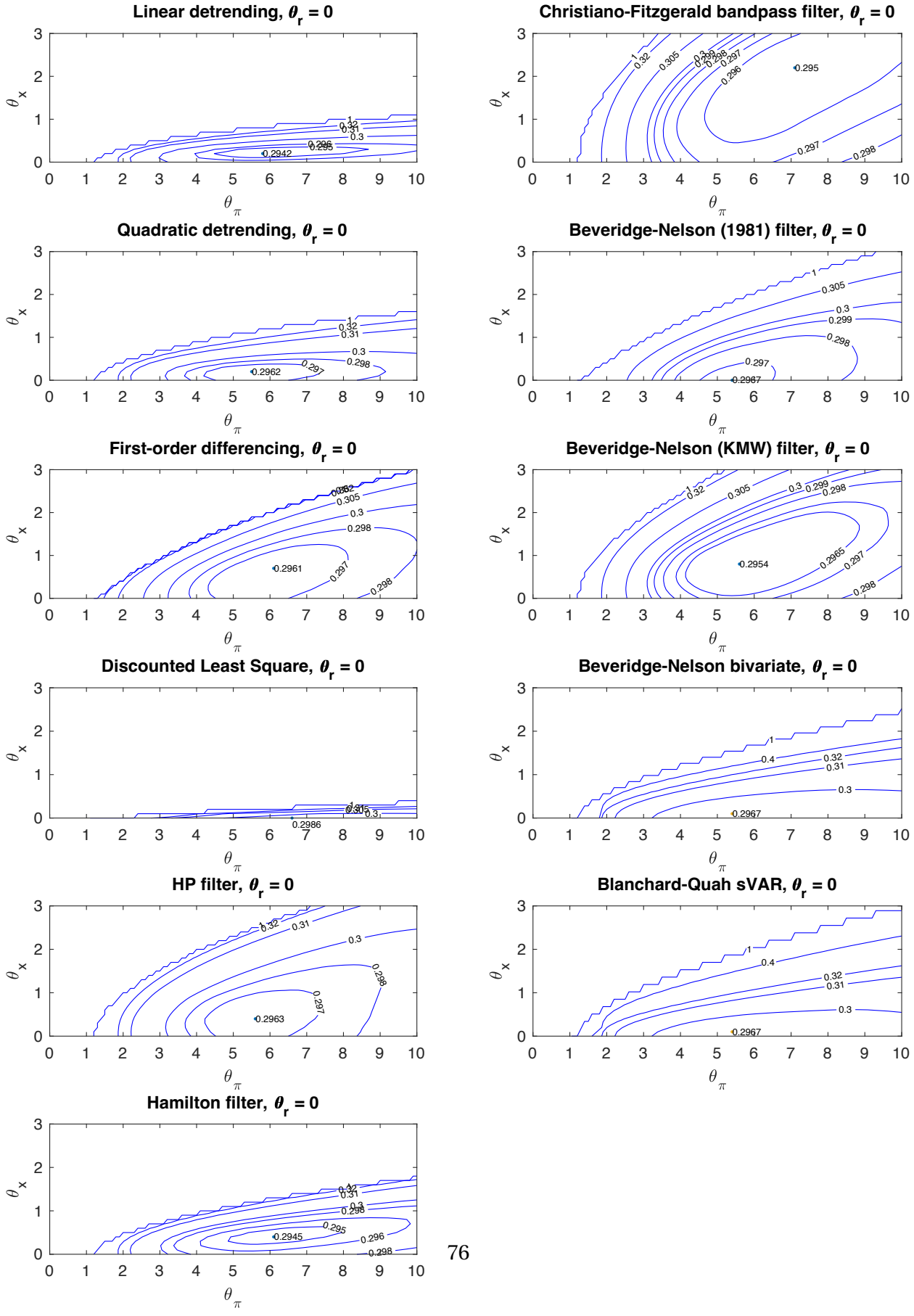


Table A.15: REVISIONS TO OUTPUT ESTIMATE WITH DETERMINISTIC TREND

Method	*Output gap estimate revision*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.566	0.712	2.912	-3.147	0.933
Beveridge-Nelson (KMW, 2016)	0.115	0.145	0.669	-0.651	0.945
HP filter	0.898	1.129	4.558	-4.931	0.971
Beveridge-Nelson (1981)	1.323	1.489	6.180	-6.718	0.844
Blanchard-Quah	1.477	2.433	9.716	-9.428	0.901
Beveridge-Nelson (bivariate)	1.493	2.509	10.541	-11.330	0.885
Hamilton	0.462	0.593	3.236	-4.520	0.855
Quadratic trend	1.848	2.325	8.821	-9.284	0.997
Linear trend	2.160	2.715	8.298	-11.106	0.999

Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method when the underlying trend is deterministic. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

process contains both level and growth shocks

$$\Delta a_t^p = g_t + e_{s,t}, \tag{A.3}$$

$$g_t = (1-\rho_g)\bar{g} + \rho_g g_{t-1} + v_{g,t}; \tag{A.4}$$

$$\Delta y_t^p = \Delta a_t^p, \tag{A.5}$$

where the disturbances $e_{s,t}$ and $v_{g,t}$ are assumed to follow Gaussian distributions and serially uncorrelated, with variance σ_s^2 and σ_v^2, respectively. $e_{s,t} \sim IIN(0,\sigma_s^2)$ is a permanent shock to the level of trend productivity and $v_{g,t} \sim IIN(0,\sigma_v^2)$ is a persistent shock to the growth rate of trend productivity. σ_s is calibrated to the same value as σ_a in the baseline. Following Stock and Watson (1998) and Edge et al. (2007), we assume a random walk in the trend growth rate ($\rho_g = 1$) in the numerical results below. In the literature, the standard deviation of the trend growth shock, σ_v, is found to be small, e.g. Stock and Watson (1998), Harvey (1985) and Benati (2007). σ_v is calibrated so that the signal-to-noise ratio equals to 0.03, as estimated in Stock and Watson (1998) using US data.

Panel B of Table A.14 summarizes the size of the stability region associated with each detrending method. The reductions for linear and polynomial (quadratic) detrending are the largest: from 16.3% to 4.5% for linear detrending and from 23.2% to 6.5% for polynomial detrending when $\theta_r = 0$, for example. For the remaining methods, the size of stability regions only shrinks slightly. The ranking of the 11 detrending methods, in terms of the size of the associated stability region, remains identical.

Table A.16: OPTIMAL POLICY AND WELFARE LOSS UNDER ALTERNATIVE TREND ASSUMPTIONS

Panel A: Deterministic trend							Panel B: Shocks to trend growth and level						
Method	*Coeff*		*SD*					*Coeff*		*SD*			
	θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)		θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)
Part 1: methods intrinsically producing smaller and less volatile $\widetilde{x}_{t-1}$													
Christiano-Fitzgerald band-pass filter													
$\theta_r = 0$	7.1	2.2	0.477	4.734	0.295	56.1	$\theta_r = 0$	7.1	0.5	0.479	4.942	0.302	52.0
$\theta_r = 0.6$	2.8	1.1	0.476	4.755	0.295	40.5	$\theta_r = 0.6$	3.1	0.4	0.476	4.991	0.301	36.6
Beveridge-Nelson (KMW)													
$\theta_r = 0$	5.6	0.8	0.478	4.739	0.295	44.7	$\theta_r = 0$	7.2	0.2	0.479	4.949	0.302	52.2
$\theta_r = 0.6$	2.6	0.6	0.476	4.784	0.295	33.6	$\theta_r = 0.6$	3.2	0.3	0.476	4.991	0.301	37.3
HP filter													
$\theta_r = 0$	5.6	0.4	0.477	4.784	0.296	43.5	$\theta_r = 0$	7.1	0.1	0.479	4.954	0.303	51.6
$\theta_r = 0.6$	2.4	0.4	0.476	4.828	0.297	30.6	$\theta_r = 0.6$	3.0	0.2	0.476	5.006	0.302	34.9
First-order differencing													
$\theta_r = 0$	6.1	0.7	0.478	4.747	0.296	50.0	$\theta_r = 0$	7.1	0.2	0.479	4.942	0.303	52.1
$\theta_r = 0.6$	2.8	0.6	0.476	4.772	0.295	38.4	$\theta_r = 0.6$	3.6	0.6	0.477	4.938	0.300	44.4
Beveridge-Nelson (1981)													
$\theta_r = 0$	5.4	0.0	0.477	4.794	0.297	41.4	$\theta_r = 0$	7.0	0.0	0.479	4.954	0.303	51.1
$\theta_r = 0.6$	2.6	0.4	0.477	4.830	0.297	34.2	$\theta_r = 0.6$	3.1	0.3	0.476	4.977	0.301	37.6
Part 2: methods intrinsically producing larger and more volatile $\widetilde{x}_{t-1}$													
Blanchard-Quah													
$\theta_r = 0$	5.4	0.1	0.477	4.794	0.297	41.5	$\theta_r = 0$	7.0	0.1	0.479	4.954	0.303	51.1
$\theta_r = 0.6$	2.3	0.0	0.476	4.885	0.298	25.6	$\theta_r = 0.6$	3.0	0.0	0.476	5.022	0.302	34.0
Beveridge-Nelson (bivariate)													
$\theta_r = 0$	5.4	0.1	0.477	4.794	0.297	41.5	$\theta_r = 0$	7.0	0.1	0.479	4.954	0.303	51.1
$\theta_r = 0.6$	2.3	0.0	0.476	4.885	0.298	25.5	$\theta_r = 0.6$	3.0	0.0	0.476	5.022	0.302	34.0
Hamilton													
$\theta_r = 0$	6.1	0.4	0.477	4.723	0.295	48.8	$\theta_r = 0$	6.9	0.1	0.479	4.937	0.302	50.6
$\theta_r = 0.6$	3.1	0.3	0.475	4.769	0.294	42.3	$\theta_r = 0.6$	3.1	0.1	0.476	4.987	0.301	36.1
Quadratic trend													
$\theta_r = 0$	5.5	0.2	0.477	4.774	0.296	43.3	$\theta_r = 0$	7.0	0.0	0.479	4.954	0.303	51.1
$\theta_r = 0.6$	2.5	0.1	0.476	4.875	0.297	29.7	$\theta_r = 0.6$	3.0	0.0	0.476	5.022	0.302	34.0
Linear trend													
$\theta_r = 0$	5.8	0.2	0.477	4.708	0.294	45.9	$\theta_r = 0$	7.7	0.1	0.479	4.941	0.302	54.8
$\theta_r = 0.6$	2.6	0.1	0.475	4.784	0.295	32.1	$\theta_r = 0.6$	3.0	0.0	0.476	5.022	0.302	34.0
Discounted Least Squares													
$\theta_r = 0$	5.4	0.0	0.477	4.794	0.297	41.5	$\theta_r = 0$	7.0	0.0	0.479	4.954	0.303	51.1
$\theta_r = 0.6$	2.3	0.0	0.476	4.885	0.298	25.5	$\theta_r = 0.6$	3.0	0.0	0.476	5.022	0.302	34.0

Notes: This table reports optimized policy coefficients and welfare loss for each detrending method under alternative trend assumptions. Panel A assumes a deterministic trend; Panel B assumes shocks to both the level and growth rate of the trend. SD denotes standard deviation from simulation.

Panel B of Table A.16 shows the optimized policy coefficients and welfare loss under this scenario. Comparing with the baseline results, the minimized welfare losses obtained here are higher for every method and the optimized policy coefficient on output gap estimates θ_x is lower.

Table A.17 reports 10-year revisions to output gap estimates associated with the optimized policy coefficients with no inertia when the true underlying trend is stochastic with shocks to both trend level and growth rates. The statistical properties of revisions to output gap estimates are similar to results in Table 3.

Table A.17: REVISIONS TO OUTPUT ESTIMATE (BOTH LEVEL AND GROWTH SHOCKS TO THE TREND)

Method	*Output gap estimate revision*				
	Mean	*SD*	*Max*	*Min*	*AR*
Christiano-Fitzgerald	0.633	0.796	3.122	-3.480	0.938
Beveridge-Nelson (KMW, 2016)	0.163	0.204	0.705	-0.847	0.974
HP filter	1.015	1.278	5.186	-5.686	0.974
Beveridge-Nelson (1981)	1.308	1.506	6.583	-8.203	0.796
Blanchard-Quah	1.489	2.175	9.624	-10.955	0.905
Beveridge-Nelson (bivariate)	1.490	2.415	12.837	-12.296	0.881
Hamilton	0.560	0.720	4.582	-4.043	0.875
Quadratic trend	2.210	2.776	9.956	-10.330	0.997
Linear trend	2.707	3.381	10.581	-11.857	0.999

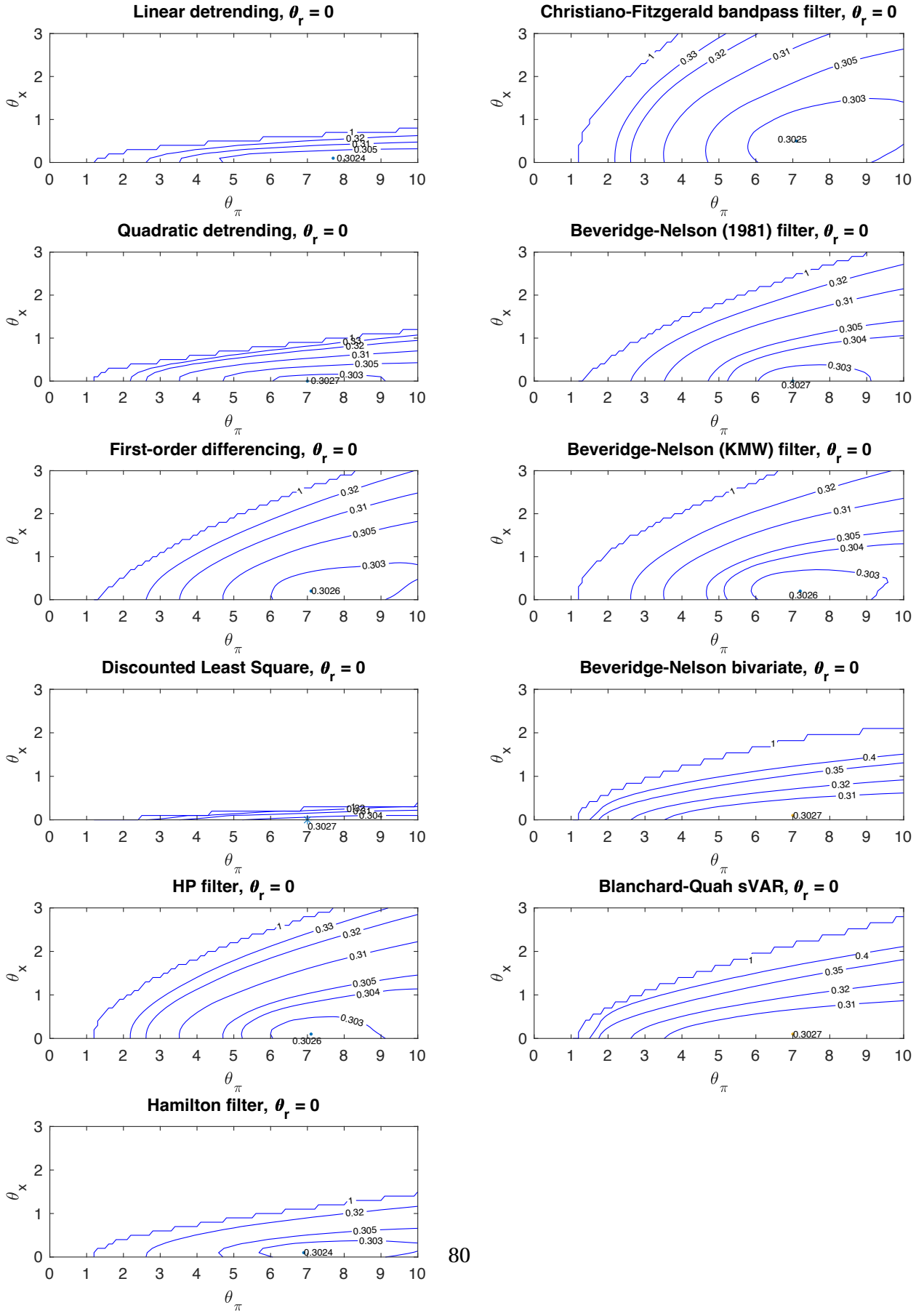
Notes: This table reports the statistical properties of 10-year revisions to output gap estimates for each detrending method when the underlying trend has both level and growth shocks. Revisions are calculated as the estimate of the current output gap made 10 years later minus the real-time estimate.

Figure A.20 plots the iso-loss contours for each detrending method under this scenario when $\theta_r = 0$. The order of the size of stability region roughly remains unchanged but the stability region shrinks, compared to the baseline results.

M.3 Robust monetary policy

Monetary policymakers may face uncertainty about the true underlying productivity trend. For instance, the underlying trend may be deterministic (equation (A.1)-(A.2)), or stochastic with only level shocks to the trend (equation (2.3)-(2.4)), or stochastic with shocks to both trend level and growth rates (Equation (A.3)-(A.5)). This section examines the robust optimal interest rate rule against such uncertainty by a Min-Max algorithm, following Orphanides and Williams (2007), given a detrending method adopted by the policymaker. The algorithm minimizes the maximum loss associated with the worst case scenario under these alternative assumptions. The robust optimal policy and associated welfare loss according to the Min-Max algorithm is

Figure A.20: ISO-LOSS CONTOURS FOR DIFFERENT DETRENDING METHODS (WITH BOTH LEVEL AND GROWTH SHOCKS TO THE TREND)



obtained in three steps. First, for each pair of policy coefficients (θ_π and θ_x), we simulate the model under the three alternative assumptions: a stochastic trend with only level shocks to the trend, the deterministic trend and a stochastic trend with shocks to both trend level and growth, respectively. Then we construct a new set of losses by finding the maximum loss for each pair of policy coefficients from the three alternative sets of simulations (the Max-step). Finally, we find the policy coefficients which minimize the losses in the previous step (the Min-step).

We find that the Min-Max monetary policy strategies are identical to the optimal policy in Panel B of Table A.16 with both level and growth shocks to the trend. This is because the welfare loss associated with this scenario is higher than that under the other two assumptions about the trend, given a choice of policy coefficients. Thus, the policymaker facing uncertainty about the underlying trend can achieve the Min-Max robust outcomes by making policy decisions based on the optimized rules as in the case with both level and growth shocks to the trend.

The robust policy exercise above provides recommendations on the conduct of policy, given a choice of detrending method. On the other hand, irrespective of the assumption on the underlying trend and the choice of detrending method, a sufficiently large response to inflation and small response to output gap estimates robustly delivers good welfare performance. This is clear from the iso-loss contours under all scenarios (with different trend assumptions and across different detrending methods).

N Different speeds of learning by households

The upper panel of Table A.18 reports the stability region for each method when the household's learning is slower, i.e. constant gains here are halved compared to the baseline. The lower panel of Table A.18 reports the results when the household's learning is faster, that is, the values of the constant gain parameter here equal 1.5 times the baseline values. Panel A and Panel B of Table A.19 report optimized monetary policy coefficients and associated welfare loss and proportions of hitting ZLB when the household is learning slower and faster, respectively. Unsurprisingly, the size of stability regions depend negatively on the gain parameters. The ranking of detrending methods does not change; this result and other main results are robust to alternative gain parameter values.

O Alternative model parameterization and belief specification

This appendix considers an alternative parameterization and households' forecast model. Households have a correctly-specified forecasting model of output, inflation and nominal interest rates.

Table A.18: THE SIZE OF STABILITY REGION FOR DETRENDING METHODS (DIFFERENT SPEEDS OF LEARNING)

Stability region (%)				
Lower gains (×0.5)	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	82.3	83.1	83.5	82.9
Beveridge-Nelson (KMW, 2016)	75.6	76.0	75.5	73.6
HP filter	71.1	71.3	70.1	68.1
First-order differencing	65.4	67.5	71.2	72.3
Beveridge-Nelson (1981)	63.8	66.0	69.4	70.7
Blanchard-Quah	67.2	63.8	60.1	55.7
Beveridge-Nelson (bivariate)	54.7	48.2	41.2	32.2
Hamilton	37.0	36.0	33.7	31.1
Quadratic trend	31.0	30.3	26.9	20.5
Linear trend	24.3	18.5	13.1	10.1
Discounted Least Squares	8.6	8.8	8.9	8.5
Higher gain (×1.5)	$\theta_r = 0$	$\theta_r = 0.2$	$\theta_r = 0.6$	$\theta_r = 0.8$
Christiano-Fitzgerald	74.9	76.0	75.7	71.8
Beveridge-Nelson (KMW, 2016)	72.2	72.8	70.6	62.1
HP filter	57.5	57.0	50.7	38.8
First-order differencing	51.5	54.7	60.0	56.6
Beveridge-Nelson (1981)	47.5	50.6	55.4	37.2
Blanchard-Quah	62.1	53.1	34.0	29.0
Beveridge-Nelson (bivariate)	47.8	36.2	28.8	19.4
Hamilton	22.6	22.2	19.8	16.0
Quadratic trend	18.7	18.8	17.1	3.8
Linear trend	11.0	10.5	8.1	1.4
Discounted Least Squares	4.2	4.2	4.0	0.5

Notes: This table reports the size of the stability region for each detrending method under different speeds of household learning. The upper panel uses gain parameters halved relative to baseline; the lower panel uses gain parameters 1.5 times the baseline. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. The ZLB is imposed.

Table A.19: OPTIMIZED POLICY COEFFICIENTS AND WELFARE LOSS: LOWER AND HIGHER GAIN PARAMETERS

Panel A: Lower gains (×0.5)							Panel B: Higher gains (×1.5)						
Method	Coeff		SD					Coeff		SD			
	θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)		θ_π	θ_x	σ_π	σ_x	*Loss*	ZLB(%)
Christiano-Fitzgerald band-pass filter													
$\theta_r = 0$	10	0.6	0.464	4.474	0.276	61.9	$\theta_r = 0$	5.4	0.8	0.494	5.308	0.329	44.4
$\theta_r = 0.6$	6.1	0.5	0.462	4.476	0.274	55.9	$\theta_r = 0.6$	2.2	0.5	0.493	5.366	0.329	29.5
Beveridge-Nelson (Kamber, Morley and Wong, 2018)													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.8	$\theta_r = 0$	5.4	0.5	0.494	5.303	0.329	44.0
$\theta_r = 0.6$	5.6	0.2	0.463	4.479	0.274	52.8	$\theta_r = 0.6$	2.2	0.4	0.493	5.338	0.329	29.6
HP filter													
$\theta_r = 0$	10	0.5	0.464	4.477	0.276	62.0	$\theta_r = 0$	5.5	0.2	0.494	5.355	0.330	43.7
$\theta_r = 0.6$	5.7	0.2	0.462	4.481	0.274	53.2	$\theta_r = 0.6$	2.0	0.2	0.494	5.420	0.332	25.0
First-order differencing													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.7	$\theta_r = 0$	5.7	0.7	0.495	5.267	0.329	49.2
$\theta_r = 0.6$	5.2	0.5	0.463	4.476	0.274	52.6	$\theta_r = 0.6$	2.8	0.7	0.494	5.262	0.327	42.3
Beveridge-Nelson (1981)													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.8	$\theta_r = 0$	5.7	0.5	0.495	5.304	0.329	47.9
$\theta_r = 0.6$	6.0	0.2	0.463	4.479	0.274	55.5	$\theta_r = 0.6$	2.4	0.5	0.494	5.307	0.328	36.3
Blanchard-Quah													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.8	$\theta_r = 0$	5.5	0.1	0.494	5.363	0.330	43.7
$\theta_r = 0.6$	5.6	0.0	0.463	4.483	0.274	52.6	$\theta_r = 0.6$	2.0	0.0	0.493	5.450	0.332	23.8
Beveridge-Nelson (bivariate)													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.8	$\theta_r = 0$	5.5	0.1	0.494	5.363	0.330	43.7
$\theta_r = 0.6$	5.6	0.0	0.463	4.483	0.274	52.6	$\theta_r = 0.6$	2.0	0.0	0.493	5.450	0.332	23.8
Hamilton													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.7	$\theta_r = 0$	5.7	0.3	0.494	5.284	0.328	47.0
$\theta_r = 0.6$	5.8	0.1	0.463	4.479	0.274	54.0	$\theta_r = 0.6$	2.1	0.1	0.492	5.363	0.329	26.4
Quadratic trend													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.7	$\theta_r = 0$	5.3	0.1	0.494	5.351	0.330	42.8
$\theta_r = 0.6$	5.6	0.0	0.463	4.483	0.274	52.6	$\theta_r = 0.6$	2.0	0.0	0.493	5.450	0.332	23.8
Linear trend													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.7	$\theta_r = 0$	5.4	0.1	0.494	5.280	0.327	43.5
$\theta_r = 0.6$	5.6	0.0	0.463	4.483	0.274	52.6	$\theta_r = 0.6$	2.6	0.1	0.493	5.362	0.329	34.4
Discounted Least Squares													
$\theta_r = 0$	10	0.0	0.464	4.478	0.276	61.7	$\theta_r = 0$	5.5	0.0	0.494	5.363	0.330	43.7
$\theta_r = 0.6$	5.6	0.0	0.463	4.483	0.274	52.6	$\theta_r = 0.6$	2.0	0.0	0.493	5.450	0.332	23.8

Notes: This table reports optimized policy coefficients and welfare loss for each detrending method under different speeds of household learning. Panel A uses gain parameters halved relative to baseline (slower learning); Panel B uses gain parameters 1.5 times the baseline (faster learning). SD denotes standard deviation from simulation.

We follow the model specification of Orphanides and Williams (2007, 2008) and convert their unemployment gap "IS" curve to an output gap "IS" curve at a quarterly frequency:

$$x_t = 0.5x^e_{t+1} + 0.5x_{t-1} - 0.4(i^e_t - \pi^e_{t+1}) + e_{x,t}, \tag{A.1}$$

$$e_{x,t} = 0.513e_{x,t-1} + v_{x,t}, \tag{A.2}$$

where $v_{x,t} \sim IIN(0, \sigma^2_x)$ and $\sigma_x = 0.13\%$. The Phillips curve remains identical to the baseline.

Unlike our baseline scenario where households are assumed to have simple, underparametrized rules, we now assume they have a forecasting model which nests the RE solution of the model without ZLB, as in, for instance, Eusepi and Preston (2011) and Kuang (2014). Let Y_t denote the 1×3 vector: $Y_t = (\Delta y_t, \Delta\pi_t, \Delta i_t)$. Let X_t be the 8×1 vector of regressors in the forecasting model: $X_t = (1, \Delta y_{t-1}, \Delta\pi_{t-1}, \Delta i_{t-1}, e_{x,t-1}, e_{\pi,t-1}, e_{i,t-1}, e_{a,t-1})$. Denote by c_t the 8×3 matrix of coefficients of the forecasting model. Households use a constant-gain least squares algorithm to recursively update their beliefs about the parameters in the forecasting model:

$$c_t = c_{t-1} + \Psi R_t^{-1} X_t (Y_t - X_t' c_{t-1}),$$
$$R_t = R_{t-1} + \Psi (X_t X_t' - R_{t-1}),$$

where Ψ is the constant gain and set to 0.002, as in e.g., Eusepi and Preston (2011). Households' beliefs are initialized from RE values. Instead of conducting a long simulation as in the baseline, we simulate the model for 160 periods with 500 repetitions. Figure A.21 plots iso-loss contours in this case (without interest rate inertia). For each choice of policy coefficients, the median loss across repetitions is used to produce the iso-loss contours.

Table A.20 reports the size of stability regions, optimized coefficients, welfare loss for each detrending method. For each OG estimation method, the size of stability region is calculated as the ratio of the area of the parameter space with a welfare loss no more than 10 to the area of the full parameter space, i.e. 30.[44] The main results are robust to this alternative model parameterization and households' forecasting model, such as the ranking of detrending methods and the optimal response to OG estimates. For instance, the bandpass and BN-KMW filters continue to have the largest stability region, while linear, quadratic, and DLS detrending have the smallest stability regions.

[44] Ten seems to be the appropriate threshold value in this case. The size of stability region will be similar when we change this threshold value to a larger (e.g., 100) or a smaller number (e.g., 2).

Figure A.21: ISO-LOSS CONTOURS FOR ALTERNATIVE MODEL WITH CORRECTLY SPECIFIED BELIEF AND $\theta_r = 0$

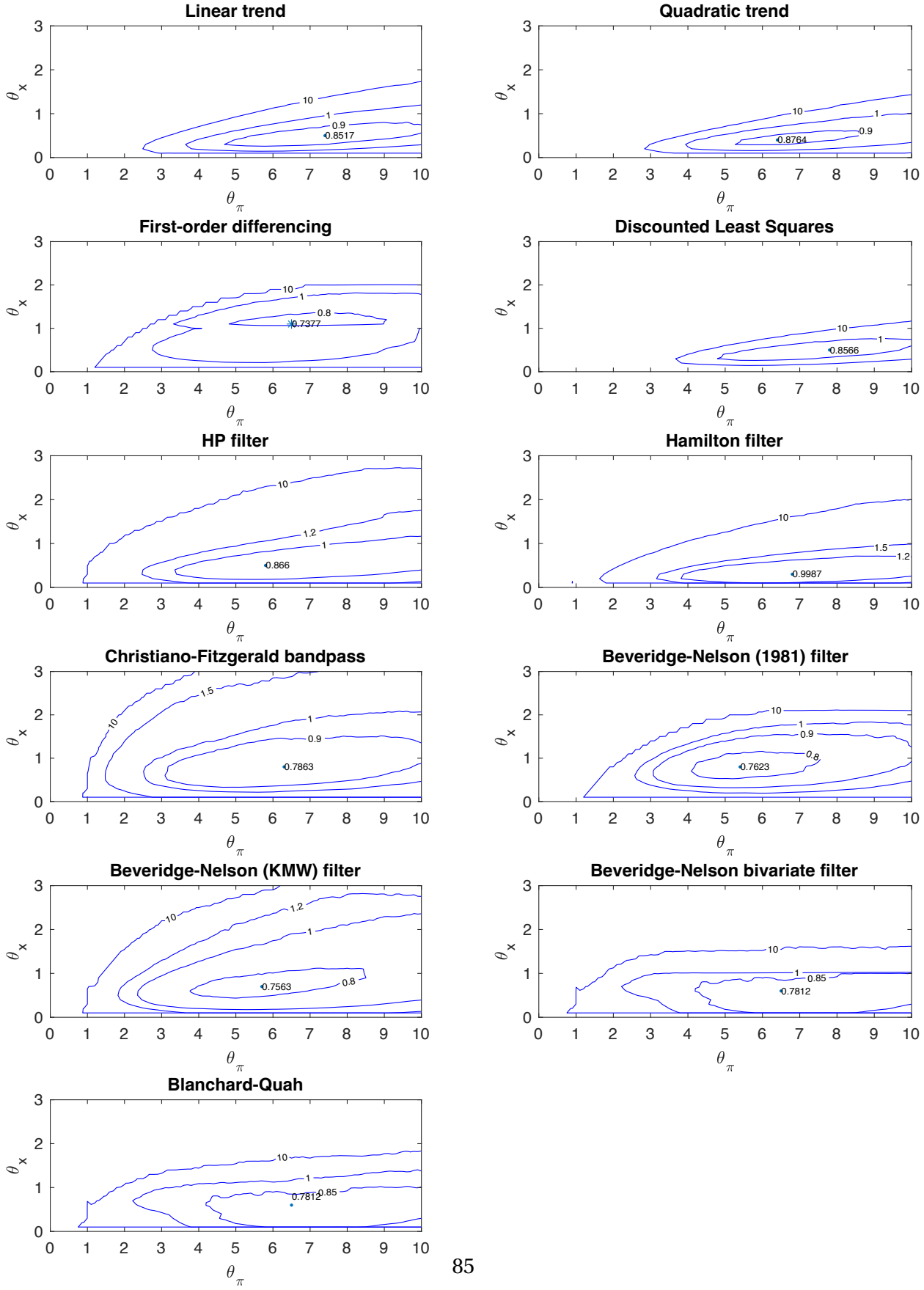


Table A.20: ALTERNATIVE MODEL AND CORRECTLY SPECIFIED BELIEF: STABILITY, OPTIMIZED POLICY COEFFICIENTS, WELFARE LOSS

Alternative model and belief				
		Coeff		
Method	Stability(%)	θ_π	θ_x	Loss
Christiano-Fitzgerald	79.5	6.3	0.8	0.786
Beveridge-Nelson (KMW)	72.9	5.7	0.7	0.756
HP filter	63.5	5.8	0.5	0.866
First-order differencing	58.8	6.5	1.1	0.738
Beveridge-Nelson (1981)	48.2	5.4	0.8	0.762
Blanchard-Quah	42.1	6.5	0.6	0.781
Beveridge-Nelson (bivariate)	41.0	6.5	0.6	0.781
Hamilton	38.1	6.8	0.3	0.999
Quadratic trend	15.7	6.4	0.4	0.876
Linear trend	25.4	7.4	0.5	0.852
Discounted Least Squares	10.5	7.8	0.5	0.857

Notes: This table reports the size of the stability region, optimized policy coefficients, and welfare loss for each detrending method under the alternative model with correctly specified household beliefs. The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 10. The ZLB is imposed.

P Infinite horizon learning model

This appendix considers a formulation motivated in Preston (2005, 2006) where he demonstrates an alternative learning approach with agents' consumption and price setting decisions depending on their forecasts into the entire infinite future. Consider the following New Keynesian model in this vein from Eusepi and Preston (2018) (except using $t-1$ dated expectation):

$$x_t = \widetilde{E}_{t-1}\left\{\sum_{T=t}^{\infty}\beta^{T-t}\left[(1-\beta)x_{T+1}-(i_T-\pi_{T+1})\right]\right\},$$
$$\pi_t = \widetilde{E}_{t-1}\left\{\sum_{T=t}^{\infty}(\alpha\beta)^{T-t}\left[\kappa\left(x_{T+1}+u_t\right)+(1-\alpha)\beta\pi_{T+1}\right]\right\}.$$

where $0<\alpha<1$ is the exogenous probability of the firms not being able to reset their price in any subsequent period, $\kappa = \frac{(1-\alpha\beta)(1-\alpha)}{\alpha}$ and u_t is an exogenous i.i.d Gaussian process reflecting variations in the elasticity of demand.[45] Following Eusepi and Preston (2018), we set α and β to 0.7 and 0.99, respectively. Households' forecasting model and belief updating remain the same as in the baseline case. Following Williams (2006), we approximate this infinite-horizon sum of

[45] Strictly speaking, the above equations are obtained from a purely forward-looking New Keynesian model, unlike the current model which has a small weight on the lagged OG. The results here are only meant to be illustrative.

expectation with a truncated long-horizon (k-period) sum of expectation with $k = 20$.

$$\begin{aligned} x_t &= \widetilde{E}_{t-1}\left\{\sum_{T=t}^{t+k-1}\beta^{T-t}\left[(1-\beta)x_{T+1}-(i_T-\pi_{T+1})\right]\right\} \\ &= (1-\beta)\left[\sum_{T=t}^{t+k-1}\beta^{T-t}\widetilde{E}_{t-1}x_{T+1}\right]-\left[\sum_{T=t}^{t+k-1}\beta^{T-t}\left(\widetilde{E}_{t-1}i_T-\widetilde{E}_{t-1}\pi_{T+1}\right)\right]. \end{aligned} \tag{A.1}$$

So

$$\begin{aligned} x_t &= \left(1-\beta\right)\left(\gamma_t^{\Delta y}-g\right)\left(2+3\beta+4\beta^2+...+(k+1)\beta^{k-1}\right) \\ &\quad +\left(1-\beta\right)\left(1+\beta+\beta^2+...+\beta^{k-1}\right)x_{t-1} \\ &\quad -\left(\gamma_t^i-\gamma_t^\pi\right)\left(1+\beta+\beta^2+...+\beta^{k-1}\right), \end{aligned} \tag{A.2}$$

and the truncated "IS" is

$$x_t = (1-\beta^k)x_{t-1}+\frac{2-\beta-(k+2)\beta^k+(k+1)\beta^{k+1}}{1-\beta}(\gamma_t^{\Delta y}-g)-\frac{1-\beta^k}{1-\beta}\left(\gamma_t^i-\gamma_t^\pi\right). \tag{A.3}$$

Similarly, the truncated Phillips curve is

$$\begin{aligned} \pi_t &= \frac{1-\alpha}{\alpha}\left(1-(\alpha\beta)^k\right)x_{t-1}+\frac{1-\alpha}{\alpha}\cdot\frac{2-\alpha\beta-(k+2)(\alpha\beta)^k+(k+1)(\alpha\beta)^{k+1}}{1-\alpha\beta}(\gamma_t^{\Delta y}-g) \\ &\quad +\frac{(1-\alpha)\beta\left(1-(\alpha\beta)^k\right)}{1-\alpha\beta}\gamma_t^\pi. \end{aligned} \tag{A.4}$$

Table A.21 reports the size of stability region and the optimal policy in the infinite horizon learning model and Figure A.22 displays the iso-loss contours with zero inertia. All the main results continue to remain robust to this alternative learning approach.

Figure A.22: ISO-LOSS CONTOURS FOR THE INFINITE-HORIZON LEARNING MODEL AND $\theta_r = 0$

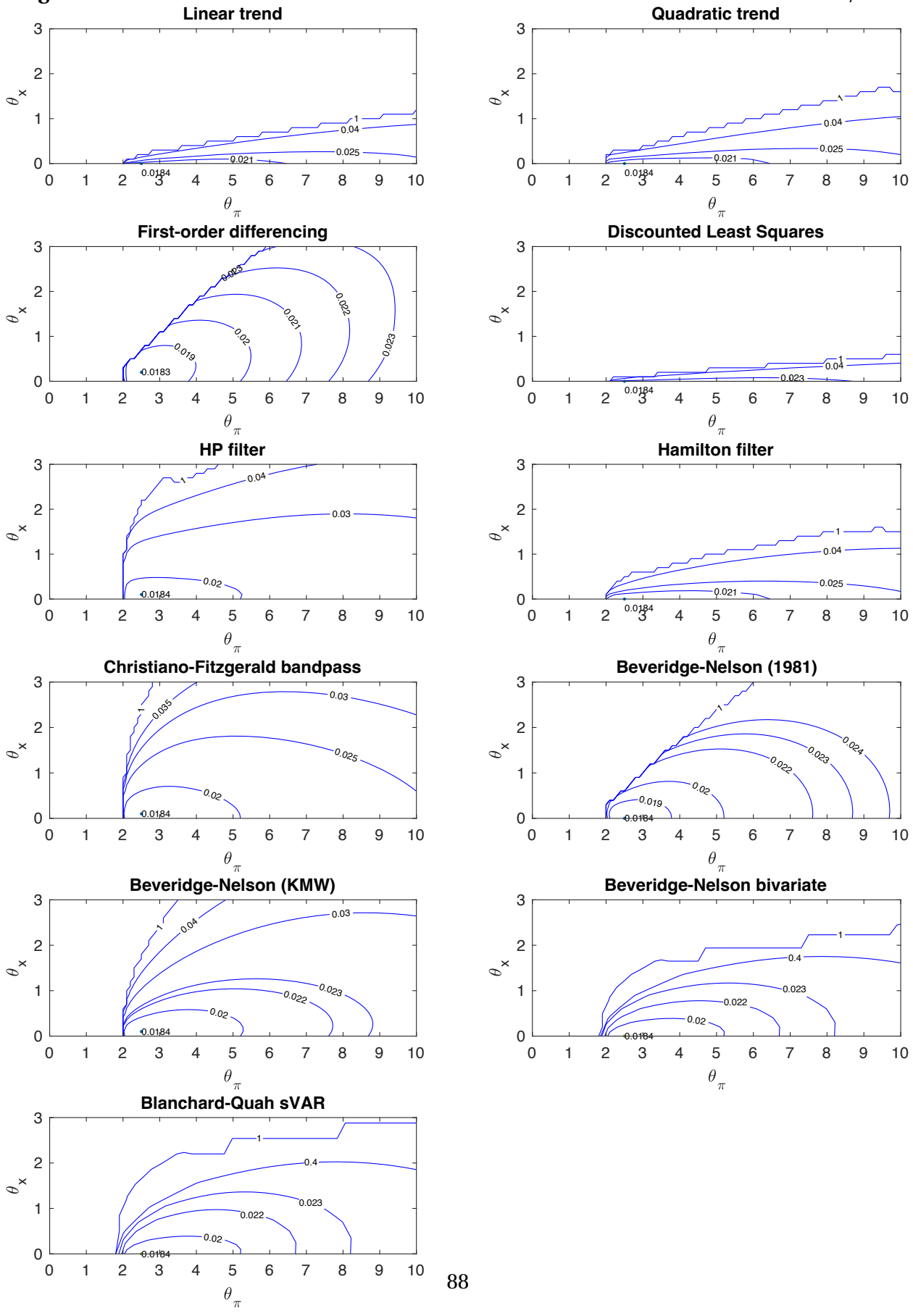


Table A.21: INFINITE-HORIZON LEARNING MODEL: STABILITY, OPTIMIZED POLICY COEFFICIENTS, WELFARE LOSS

Alternative model and belief		Coeff		
Method	Stability(%)	θ_π	θ_x	Loss
Christiano-Fitzgerald	77.5	2.5	0.1	0.018
Beveridge-Nelson (KMW)	75.3	2.5	0.1	0.018
HP filter	75.1	2.5	0.1	0.018
First-order differencing	63.4	2.5	0.2	0.018
Beveridge-Nelson (1981)	62.4	2.5	0.0	0.018
Blanchard-Quah	65.2	2.5	0.0	0.018
Beveridge-Nelson (bivariate)	50.9	2.5	0.0	0.018
Hamilton	29.3	2.5	0.0	0.018
Quadratic trend	26.3	2.5	0.0	0.018
Linear trend	17.7	2.5	0.0	0.018
Discounted Least Squares	8.6	2.5	0.0	0.018

Notes: This table reports the size of the stability region, optimized policy coefficients, and welfare loss for each detrending method in the infinite-horizon learning model with no interest rate inertia ($\theta_r = 0$). The stability region is defined as the share of the policy parameter space ($\theta_x \in [0,3]$, $\theta_\pi \in [0,10]$) for which the welfare loss is less than 1. The ZLB is imposed.

Additional References for the Appendix

Baxter, M. and King, R. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593.

Benati, L. (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics & Control*, 31:2847–2877.

Edge, R., Laubach, T., and Williams, J. (2007). Learning and shifts in long-run productivity growth. *Journal of Monetary Economics*, 54(8):2421–2438.

Evans, G. W., Honkapohja, S., and Mitra, K. (2022). Expectations, stagnation, and fiscal policy: A nonlinear analysis. *International Economic Review*, 63(3):1397–1425.

Harvey, A. and Trimbur, T. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85(2):244–255.

Harvey, C. (1985). Trends and cycles in macroeconomic time series. *Journal of Business and Economic Statistics*, 3(3):216–227.

Harvey, C. and Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–247.

Kamber, G., Morley, J., and Wong, B. (2016). Intuitive and reliable estimates of the output gap from a Beveridge-Nelson filter. BIS Working Papers 584, Bank for International Settlements.

King, G. and Rebelo, S. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17:207–231.

Lansing, K. (2002). Learning about a shift in trend output: Implications for monetary policy and inflation. Working paper.

L'Huillier, J. and Schoenle, R. (2021). Raising the inflation target: What are the effective gains in policy room? Working paper.

Powell, J. (2020). New economic challenges and the Fed's monetary policy review. Speech by Chair Powell at the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.

Stock, J. and Watson, W. (1998). Median-unbiased estimation of coefficient variance in a time-varying parameter model. *Journal of the American Statistical Association*, 93(441):349–358.