

Monetary Policy under Data Uncertainty: Interest-Rate Smoothing from a Cross-Country Perspective*

Saiah Lee[†]

University of Wisconsin–Madison

February 15, 2019

[\[Click Here for the Most Recent Version\]](#)

Cross-country estimates of Taylor rules suggest that higher data uncertainty is associated with a more inertial behavior of interest rates. Data uncertainty is measured by the volatility of differences between real-time data and their revisions. Using a simple structural model with Kalman filter learning, I replicate the cross-country pattern of the inertial behavior. More inertial behavior results not because central banks gradually adjust interest rates in the face of data uncertainty, but because the central banks' inference about the true data is correlated with past interest rates. Thus, I endogenize inertial behavior of interest rates as resulting in part from the learning process.

Keywords: Monetary Policy, Data Uncertainty, Interest-Rate Smoothing, Learning

JEL Classification: D81, D83, E52, E58

*I deeply thank Kenneth West, Simeon Alder, Charles Engel, Bruce Hansen, Rasmus Lentz, and Noah Williams for their helpful guidance and advices. I am indebted to Enghin Atalay, Menzie Chinn, Keshav Dogra, Steven Durlauf, Oleg Itskhoki, Nobuhiro Kiyotaki, Fabrizio Perri, Daniel Quint, Jeffrey Smith, Thomas Winberry, Jing Cynthia Wu, and participants at the Job Market Mini-Conference, the Juli Plant Grainger Macroeconomics Workshop, and the EGSC 2018 at Washington University in St. Louis for their useful suggestions and warm encouragement. I also thank Joel Kaiyuan Han, Junhyong Kim, Soyeon Annie Lee, Zehao Li, Kurt Lunsford, Arik Roginsky, Kegen Teng Kok Tan, Woan Foong Wong, and Steve Pak Yeung Wu for their useful discussions. The most recent version of this paper is available using the link: <http://bit.ly/SaiahLee.JMP>. All errors are my own.

[†]Ph.D. Candidate, University of Wisconsin–Madison, Department of Economics, 1180 Observatory Drive, Madison, WI 53706-1393, USA. Homepage: <http://saiahlee.com/>, Email: saiah.lee@wisc.edu, Phone: +1-608-622-0344.

“As a general rule, the Federal Reserve tends to adjust interest rates incrementally, in a series of small or moderate steps in the same direction. ... Relatively gradual policy adjustment produces better results in an uncertain economic environment.”

– Ben S. Bernanke, May 20, 2004.

“Absence of intrinsic inertia appears in accord with the views of many central bankers, who often note that future policy actions will largely be contingent on incoming data and future changes in the economic outlook.”

– Glenn D. Rudebusch, 2006.

1 Introduction

This paper examines differences in monetary policy across countries with respect to the extent to which their interest rates are slow to adjust. This is related to the quality of the data on aggregate output and the inflation rate that the country produces. We demonstrate that countries with more data uncertainty are slower to adjust their interest rate, and this is largely explained by the central banks' learning process. The novelty of this paper is that the central bank is considered as an active learner: the central bank observes data with noise and makes inferences about the true data before making policy decisions.

A key question in monetary policy is how responsive central banks are to current information about the level of GDP and inflation rate when they are determining their policy interest rate. The standard theories say that a central bank that is choosing its interest rate in a forward-looking manner should not place any weight on past interest rate when it is deciding what its current interest rate should be. However, we observe in all countries that there is at least some sluggishness to the interest rate. Moreover, the sluggishness varies considerably across countries. The sluggishness is measured using Taylor rule equation, which prescribes short-term interest rate based on inflation and the output gap. For each country, policy interest rate is regressed on inflation rate, output gap, and lagged interest rate, and the estimated coefficient on lagged interest rate indicates the sluggishness.

This paper explains both the sluggishness and the heterogeneity, and that has to do with the fact that data as its disseminated is imperfect. The data uncertainty is measured by the standard deviation of differences between real-time data and their revisions after a year. We find that countries with more data uncertainty tend to have more sluggishness in the interest rate. For example, Nigeria has greater data uncertainty and more sluggish interest rate than Canada. This is not only for Canada and Nigeria, but a general data pattern among

40 countries in the sample.¹

To explain this pattern, we build a simple New Keynesian model and allow central bank to learn about the true data from noisy observations using Kalman filter. The model consists of three components: Phillips curve, IS curve, and a monetary policy rule. Central bank's objective is to minimize a loss function, which is sum of variances in inflation rate, output gap, and interest-rate shocks. Based on all available information, central bank forms belief about the true data and chooses its responds to expected inflation rate and output gap and also decides how responsive it will be by choosing the weight on lagged interest rate.

The noisy data are defined as the true data plus noise, where the noise components follow MA(1) process. Noise tends to be persistent over time in the data, and we use MA(1) to pick up the persistence. In each period, central bank observes the noisy data and decides how much weight should be placed on the new noisy information versus its past belief and past information. The weighted sum becomes the central bank's inference about the true data, and this is the standard Kalman filter learning with MA(1) noise. The assigned weight on new noisy information is called "Kalman gain," and this decreases as the data become noisier. Based on the inferences, central bank chooses the optimal responses to inflation rate and output gap and decides the optimal weight on lagged interest rate within the monetary policy rule.

We find an interesting result from this model. As data become noisier, central bank's optimal weight on the lagged interest rate becomes smaller. This is because of the following reasons. First, the learning process effectively filters out the noise in the data. Second, the noise tends to be persistent, and the learning process helps getting the additional information from the predictable portion of the noise process.

This is not contradicting the empirical cross-country finding. The data pattern is not representative of the degree of gradual adjustment but rather is an artifact of what we as the econometrician observe from ex-post data. The model is simulated hundred thousand times

¹In accordance with conventional belief, the cross-country Taylor rule estimates suggest that higher data uncertainty is associated with more inertial behavior of interest rates. This is often referred to as *intrinsic interest-rate smoothing* in the literature: the higher the data uncertainty, the greater the reduced-form estimate on lagged interest rate. I use the terms, the sluggishness to the interest rate, the inertial behavior of interest rates, and interest-rate smoothing, interchangeably since both terms are originated from reduced-form Taylor rule estimations in the literature. Another term, gradual adjustment, is defined differently from those two, and it is the incremental adjust of interest rates by central banks. There has not been clear distinctions among the terms in the literature. I argue that interest-rate smoothing is mostly a consequence of the *learning process*, rather than central banks' *gradual adjustment* of interest rates. Central banks' real-time beliefs about true data are indistinguishable from the gradual adjustment in ex-post data, both of which are largely picked up by the Taylor rule's reduced-form estimate on the lagged interest rate.

and a series of ex-post data for each level of data uncertainty are generated. We run the same reduced-form Taylor rule regression for each level of data uncertainty and demonstrate that the coefficient on lagged interest rate increases in data uncertainty. The key explanation is that the learning component is missing in the reduced-form Taylor rule estimation on the ex-post data. The coefficient on lagged interest rate is overestimated because the central bank's belief is not included in the regressions. This is a typical omitted variable bias, and this bias increases in data uncertainty. This paper explains the sluggishness in policy interest rate with central banks' learning process when they face measurement error in the real-time data.

Contribution of the paper

This study distinguishes between the sluggishness and central banks' gradual adjustment and demonstrates that the two may not move together. Whereas the conventional view in the literature considers the sluggishness evidence of central banks' gradual adjustment ([Bernanke, 2004](#)), I show the gradual adjustment to be directly induced by central banks' desire to avoid interest-rate surprises, that is, the variance of changes in interest rates in the loss function.

The paper endogenizes the sluggishness as resulting in part from central banks' learning process. Reduced-form Taylor rule estimators in the literature often include lagged interest rate as an independent variable, which substantially increases explanatory power. Including lagged interest rate is commonly justified by the assumption that central banks partially adjust the interest rate, which implies the gradual adjustment as the key explanation for the sluggishness. I show the sluggishness to exist and increase in the level of data uncertainty even if central banks' gradual adjustment of interest rates is muted by excluding the variance of changes in interest rates from the loss function. The sluggishness can be fully endogenized in the model by the learning process to which much of it accrues and the gradual adjustment of interest rates.

Literature review

This paper is closely related to the literature that studies the effect of data uncertainty on monetary policy. For example, [Rudebusch \(2001\)](#) and [Orphanides \(2003b\)](#) find that noisy economic data may lead to cautious and timid responses of policymakers. However, in empirical studies, as noted by [Rudebusch \(2001, 2006\)](#), such inducement toward timidity

appears fairly modest. These studies focus on the effect of existence versus non-existence of data uncertainty, and all of them consider only the US economy. Because that the United States has a relatively low level of data uncertainty, the scope of previous studies is quite limited.² In contrast, I consider the cross-country variation in *magnitudes* of data uncertainty and study policymakers' optimal responses.

There are three strands of literature rationalizing interest-rate smoothing: reducing interest-rate volatility, exploiting the expectation channel for monetary policy, and responding optimally to data and model uncertainty. The first strand of the literature emphasizes the costs and benefits of interest-rate smoothing arising from its effects on financial stability (Cukierman, 1991; Rudebusch and Svensson, 1999; Stein and Sunderam, 2015). The second strand of the literature analyzes the benefits of interest-rate smoothing coming from its ability to steer private-sector expectations by inducing history dependency in the policy rate (Levin, Wieland and Williams, 2003; Rotemberg and Woodford, 1999; Woodford, 1999, 2003). The last strand of the literature explores the benefit of interest-rate smoothing arising from its ability to better manage uncertainties about data, model parameters, or the structure of the economy faced by the central bank (Brainard, 1967; Milani, 2007; Sack, 1998, 2000; Rudebusch, 2001; Söderström, 2002; Orphanides, 2003a).³

None of these papers shows a clear distinction between interest-rate smoothing and central banks' gradual interest-rate adjustment nor decreasing in the degree of gradual adjustment in the face of data uncertainty. In contrast, I decompose interest-rate smoothing into gradual interest-rate adjustment, which is caused by central banks' motive to avoid interest-rate surprises, and component in the learning process correlated with past interest rates.

Policy implications

We can think about how central banks can improve their monetary policy under data uncertainty. There are two approaches we can think of: improving learning ability and improving data quality. This research provides a framework that helps analyzing these competing options. We consider two benchmark cases in addition to the learning policy. First, central bank always observes perfect information. Second, central bank observes noisy data and naively take the face value without making any inference. We compare the welfare loss under the learning policy and the naive policy at different levels of data uncertainty and

²Cross-country variation in data uncertainty is described in Figure 2 (a).

³See Coibion and Gorodnichenko (2012) for an overview of recent literature.

demonstrate that learning policy is always better than the naive policy. The gain from learning is increasing and convex in data uncertainty. Given the cost information of each option, we can conduct a cost-benefit analysis and decide which option is more cost-effective.

The remainder of the paper proceeds as follows. Section 2 provides cross-country comparison of monetary policies under data uncertainty. Section 3 introduces the Rudebusch and Svensson (1999) model with learning. Section 4 presents the model estimation along with reduced-form estimation on simulated data. Section 4.5 discusses policy implications of the study. The last section concludes.

2 Cross-Country Observation

The conventional belief — a positive relationship between data uncertainty and interest-rate smoothing — is shown as a cross-country scatter plot in figure 2 (b).⁴ Data uncertainty ($\sigma_{n\pi} + \sigma_{ny}$) is measured by the linear combination of volatilities in the differences between real-time data and their revisions, and interest-rate smoothing ($\tilde{\rho}$) is measured by the estimated coefficients on lagged interest rate in Taylor rule. Note that countries with high data uncertainty tend to have high weight on the lagged interest rate. The measurements and estimation methods are described as the following.

2.1 Interest-rate smoothing

2.1.1 Partial-adjustment Taylor rule

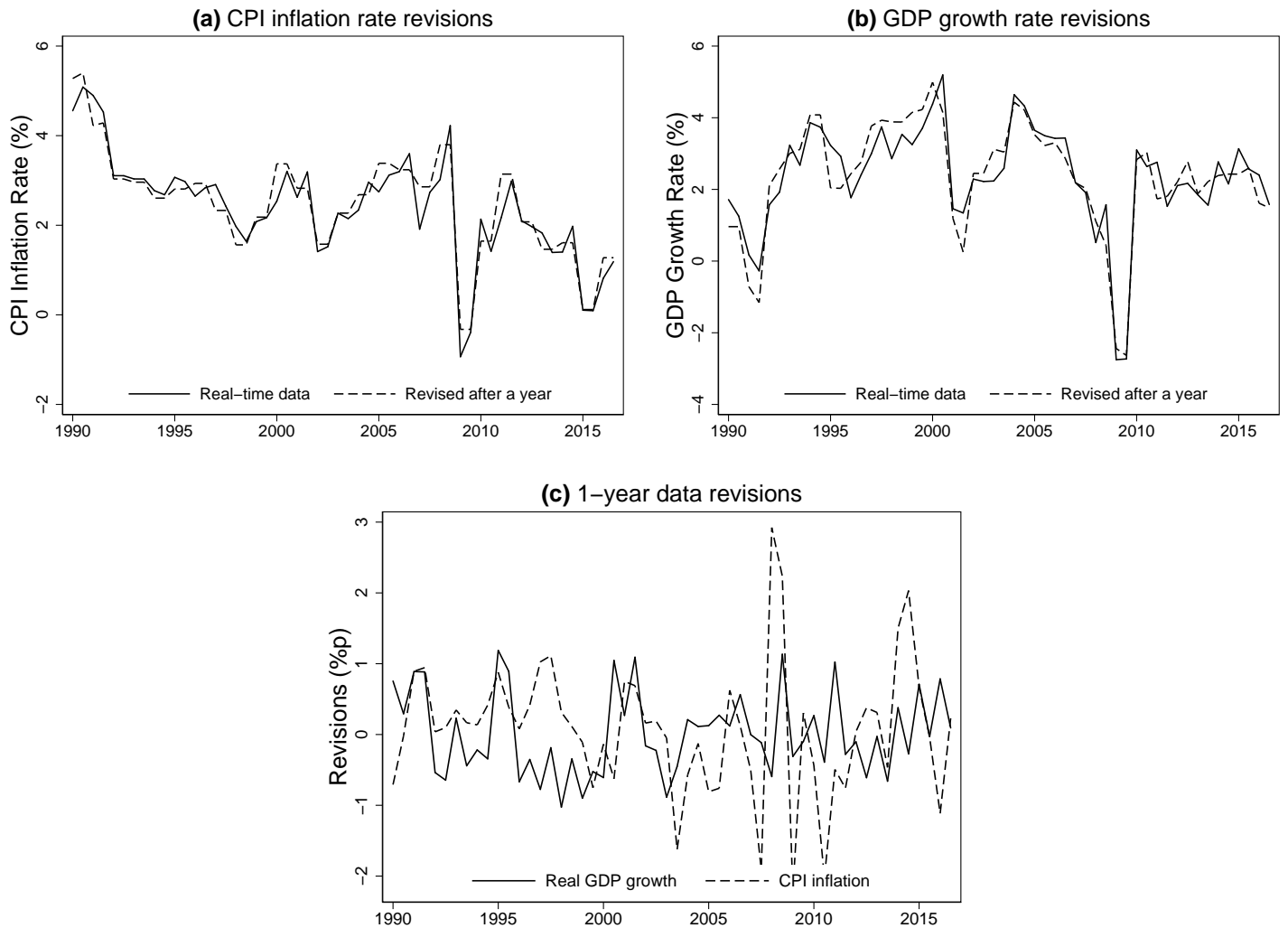
I assume that within each operating period the central bank has a target for the nominal short-term interest rate, i_t^* , that is based on the state of the economy. In the baseline case, I assume that the target depends on both expected inflation and output.

$$i_t^* = \bar{i} + g_\pi(E[\pi_{t+n}|\Omega_t] - \pi^*) + g_y(E[x_t|\Omega_t] - x_t^*) \quad (1)$$

where \bar{i} is the long-run equilibrium nominal rate, π_{t+n} is the rate of inflation between periods t and $t + n$, x_t is real output, π^* is target inflation, x^* is potential output, and Ω_t is

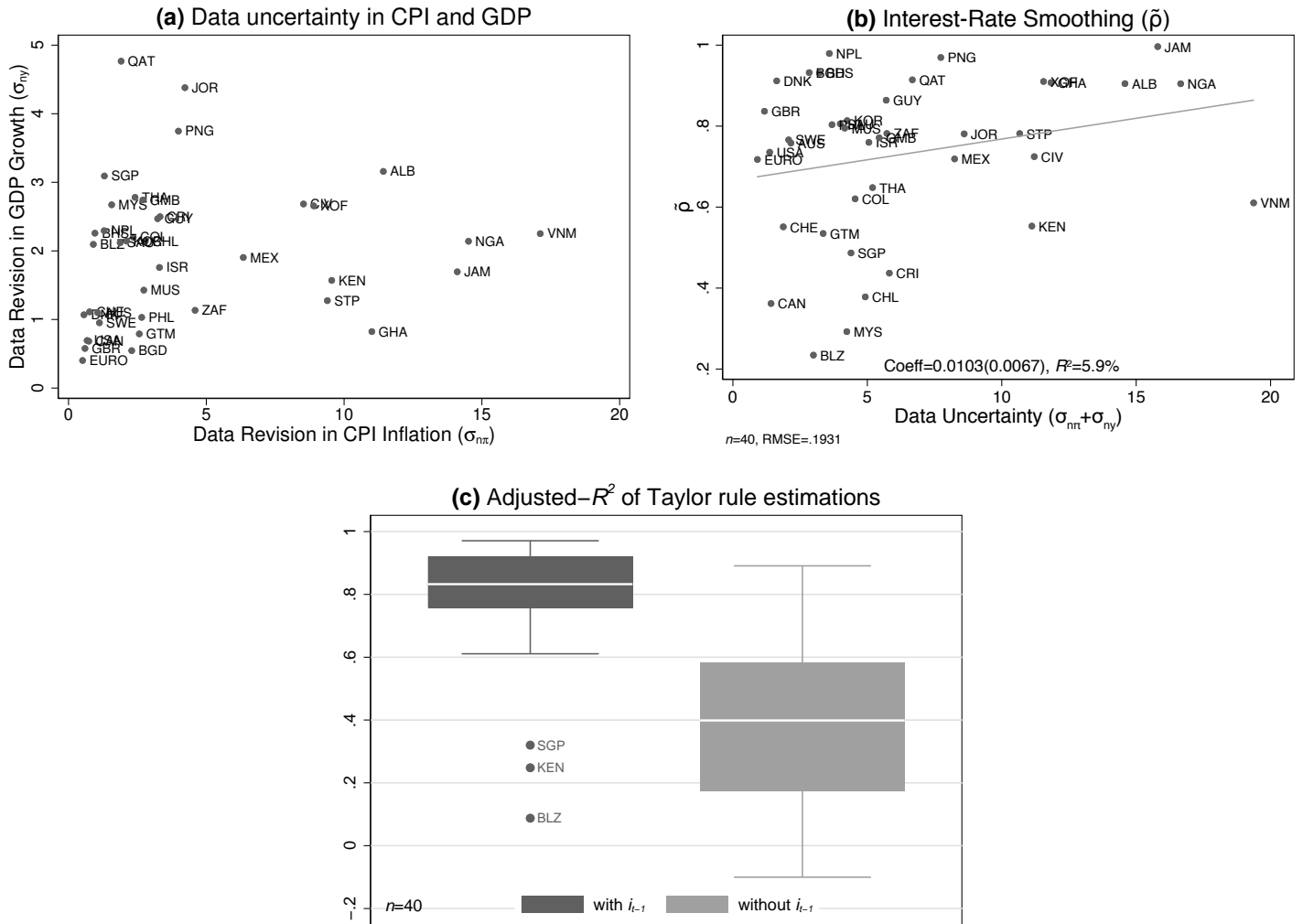
⁴The forty sample countries include: Albania, Australia, Bahamas, Bangladesh, Belize, Canada, Chile, Colombia, Costa Rica, Côte d'Ivoire, Denmark, Euro area, Gambia, Ghana, Guatemala, Guyana, Israel, Jamaica, Jordan, Kenya, Korea, Malaysia, Mali, Mauritius, Mexico, Nepal, Nigeria, Papua New Guinea, Philippines, Qatar, São Tomé and Príncipe, Saudi Arabia, Singapore, South Africa, Sweden, Switzerland, Thailand, United Kingdom, United States, and Vietnam.

Figure 1: CPI and GDP Data Revisions in the US



Note: Historical data on CPI inflation and GDP growth rate from IMF's World Economic Outlook (WEO) are used. Panel (a) and (b) show real-time data and revised data after a year in percentage, and panel (c) shows the simple differences between the real-time data and revised data after a year ($revisions_t$ (%p) = $revised\ data_t$ (%) - $real-time\ data_t$ (%)). The means of revisions in CPI inflation and real GDP growth are close to zero.

Figure 2: Data uncertainty and cross-country comparison of interest-rate smoothing ($\tilde{\rho}$)



Note: Panel (a) presents the cross-country variation of data uncertainty in CPI and GDP data. Partial-adjustment Taylor rule, $i_t = (1 - \tilde{\rho})[k + \tilde{g}_\pi E_t \pi_{t+1} + \tilde{g}_y y_t] + \tilde{\rho} i_{t-1} + \varepsilon_t$, is estimated for each country and panel (a) reports $\tilde{\rho}$ for the 40 sample countries. Panel (c) reports adjusted- R^2 s of the 40 estimations, including and excluding lagged interest rate in the Taylor rule.

the central bank's information set at time t .

I assume that the monetary policy-related interest rate partially adjusts to target, as follows:

$$i_t = (1 - \rho)i_t^* + \rho i_{t-1} + \eta_t \quad (2)$$

where the parameter $\rho \in [0, 1]$ captures the degree of interest-rate smoothing, and the exogenous random shock to the interest rate, η_t , is i.i.d. Defining $k \equiv \bar{i} - g_\pi \pi^*$ and $y_t \equiv x_t - x_t^*$, equation (1) becomes

$$i_t^* = k + g_\pi E[\pi_{t+n} | \Omega_t] + g_y E[y_t | \Omega_t] \quad (3)$$

Combining equation (2) and (3), we get

$$i_t = (1 - \rho)[k + g_\pi E[\pi_{t+n} | \Omega_t] + g_y E[y_t | \Omega_t]] + \rho i_{t-1} + \eta_t \quad (4)$$

Rewriting the policy rule in terms of realized variables at the current period and setting $n = 1$, we have

$$i_t = (1 - \rho)[k + g_\pi E_t \pi_{t+1} + g_y y_t] + \rho i_{t-1} + \varepsilon_t \quad (5)$$

where the error term ε_t is a linear combination of the forecast errors of inflation and output and the exogenous disturbance η_t . Parameters in equation (5) is estimated in each country, and the details are described below.

2.1.2 Estimation of the rule parameters

Parameters in the Taylor rule (equation (5)) are estimated for each country with least squares using quarterly data on monetary policy related interest rates, inflation rates, and output gap from 1990 to 2008. The data are drawn from the IMF's International Financial Statistics (IFS) and Thomson Reuters' Datastream. The output gap is defined by the percentage difference between real GDP and estimated potential GDP, and the potential GDP is measured by the quadratic trend of real GDP.

Rewriting equation (5) with country subscript c ,

$$i_{t,c} = (1 - \tilde{\rho}_c)[k_c + \widetilde{g_{\pi,c}} E_t \pi_{t+1,c} + \widetilde{g_{y,c}} y_{t,c}] + \tilde{\rho}_c i_{t-1,c} + \varepsilon_{t,c} \quad (6)$$

where $i_{t,c}$ is the monetary policy-related (nominal) interest rate, $\pi_{t+1,c}$ is the rate of inflation between periods t and $t + 1$, and $y_{t,c}$ is output gap. Monetary policy responses to inflation

rate and output gap are captured by the rule parameters $\widetilde{g}_{\pi,c}$ and $\widetilde{g}_{y,c}$, respectively, and k_c captures country-specific equilibrium real interest rate and target inflation together. The expected inflation rate, $E_t\pi_{t+1}$, is measured by the four quarter average inflation rate in percent following Rudebusch (2001) (i.e., $E_t\pi_{t+1} = \frac{1}{4} \sum_{j=0}^3 \pi_{t-j}$).⁵ Note that the expression, $[k + g_\pi E_t\pi_{t+1} + g_y y_t]$, represents the Taylor rule suggested optimal monetary policy. The parameter of interest, $\widetilde{\rho}_c \in [0, 1]$, captures the degree of interest-rate smoothing in country c . Estimated $\widetilde{\rho}$ for the sample countries are reported in figure 2 (b), and adjusted- R^2 for the regressions are reported as box plots in figure 2 (c).

2.2 Data uncertainty

Data uncertainty is measured by the volatility of the discrepancies between the noisy data and the true data, where the initially released real-time data and revised data after a year are taken as proxies for the noisy and the true data, respectively. It is formally defined as $(\sigma_{n\pi} + \sigma_{ny})$ where $\sigma_{n\pi}$ and σ_{ny} are standard errors of the discrepancies in inflation rate and GDP gap, respectively. I estimate $\sigma_{n\pi}$ and σ_{ny} using historical data from the IMF's World Economic Outlook (WEO), and the sample period is from 1990 to 2008.⁶ The detailed description of estimating $\sigma_{n\pi}$ and σ_{ny} is the following.

2.2.1 Measuring data uncertainty

The real-time noisy indicators on inflation rate, π_t^n , and output gap, y_t^n , are defined as:

$$\pi_t^n = \pi_t + n_t^\pi \quad (7)$$

$$y_t^n = y_t + n_t^y \quad (8)$$

where π_t and y_t are true inflation rate and output gap, and n_t^π and n_t^y are the contemporaneous measurement errors that plague the policymaker in real time, with standard errors $\sigma_{n\pi}$

⁵Alternative measures for $E_t\pi_{t+1}$, for example, linear trend forecasting, can be considered.

⁶The WEO data are released twice a year, in the spring and the fall. The database consists of macroeconomic data and forecasts submitted by country teams and vetted by the IMF's Research Department for both internal and multilateral consistency. The spring WEO is released in May up through 2001 and in April thereafter; the fall version is typically released in October, and occasionally in September. Historical WEO data are publicly available for the period from 1990 to 2017. I collect real GDP growth and CPI inflation data from this database, and each report includes revised estimates for the past years, real-time estimates for the current year, and forecasts based on the current information. An example of the historical WEO dataset is presented in table 1 in the appendix.

and σ_{ny} , respectively, and uncorrelated with π_t and y_t .

Note that the noise in the data appears to be quite persistent over time, as shown in figure 1. In order to capture this persistence, the noises in inflation rate and GDP gap, n_t^π and n_t^y , are modeled as first-order moving average (MA(1)) processes:

$$n_t^\pi = \epsilon_t^\pi + \theta^\pi \epsilon_{t-1}^\pi \quad (9)$$

$$n_t^y = \epsilon_t^y + \theta^y \epsilon_{t-1}^y \quad (10)$$

where ϵ_t^π and ϵ_t^y are normally distributed with mean zero and variance $\sigma_{\epsilon^\pi}^2$ and $\sigma_{\epsilon^y}^2$, respectively.⁷

I measure each country's data uncertainty using linear combinations of $\sigma_{n\pi}$ and σ_{ny} , keeping the ratio between the two. One important econometric issue is that the parameters in Taylor rule (equation (5)) are estimated using quarterly data, while the sources of data for estimating data uncertainty (historical WEO issues) are released only twice a year, second and fourth quarter. Moreover, the WEO issues are not exactly semiannual because each issue reports real GDP growth and CPI inflation for the current entire year, not for the current half year.

In order to resolve these issues, I introduce the following econometric technique. Let $\pi_{t,0}$ be the first released estimate of CPI inflation for the year of t and $\pi_{t,1}$, $\pi_{t,2}$ and $\pi_{t,3}$ be the following semi-annual revisions of it. $\pi_{t,0}$ releases in a second quarter, and $\pi_{t,1}$ releases in the fourth quarter in the same year. $\pi_{t,2}$ and $\pi_{t,3}$ release in the second and the fourth quarter of the following year. Assume that it takes four quarters to finalize the data since their first releases, then $\pi_{t,0} \sim \pi_{t,3}$ can be expressed as the following:⁸

$$\begin{aligned} \text{Real-time data: } \pi_{t,0} &= \frac{1}{4} [\pi_{t,q1}^n + \pi_{t,q2}^n + E_{q2}[\pi_{t,q3} + \pi_{t,q4}]] \\ \text{First revision: } \pi_{t,1} &= \frac{1}{4} [\pi_{t,q1}^n + \pi_{t,q2}^n + \pi_{t,q3}^n + \pi_{t,q4}^n] \\ \text{Second revision: } \pi_{t,2} &= \frac{1}{4} [\pi_{t,q1} + \pi_{t,q2} + \pi_{t,q3}^n + \pi_{t,q4}^n] \\ \text{Third revision: } \pi_{t,3} &= \frac{1}{4} [\pi_{t,q1} + \pi_{t,q2} + \pi_{t,q3} + \pi_{t,q4}] \end{aligned} \quad (11)$$

⁷The assumption of the persistent noise in the real-time data accords with previous studies. See [Orphanides \(2001\)](#) for a detailed argument. [Rudebusch \(2001\)](#) uses AR(1) for modeling the noise in GDP gap and MA(3) for inflation rate. [Onatski and Williams \(2003\)](#) use AR(1) for both. I use MA(1) for both, and this assumption is useful under the structure of data releases as described in the next page.

⁸On average 79% of inflation and GDP data revisions happen within four quarters from the initial release in the United States and 73% in all sample countries.

where $\pi_{t,q1} \sim \pi_{t,q4}$ are true CPI inflation in each quarter, and $\pi_{t,q1}^n \sim \pi_{t,q4}^n$ are noisy indicators of them. All variables above are annualized. $\pi_{t,0}$ consists of two noisy indicators for the first and second quarters and two forecasts for the third and fourth quarters, and $\pi_{t,1}$ consists of noisy indicators for all four quarters. $\pi_{t,2}$ consists of two true inflation rates for the first and the second quarters, because it has been four quarters since $\pi_{t,0}$ was initially released, and two noisy indicators for the third and the fourth quarters. $\pi_{t,3}$ consists of true inflation rates for all four quarters.⁹ From equation (7) and (11), note that

$$\begin{aligned}\pi_{t,1} - \pi_{t,2} &= \frac{1}{4} [n_{t,q1}^\pi + n_{t,q2}^\pi] \\ \pi_{t,2} - \pi_{t,3} &= \frac{1}{4} [n_{t,q3}^\pi + n_{t,q4}^\pi]\end{aligned}\tag{12}$$

Since n_t^π is assumed to follow MA(1), $\{\theta^\pi, \sigma_{\epsilon^\pi}^2\}$ in equation (9) can be estimated with observations $(n_{t,q1}^\pi + n_{t,q2}^\pi)$ and $(n_{t,q3}^\pi + n_{t,q4}^\pi)$ by matching the moments:

$$Var(\pi_{t,1} - \pi_{t,2}) = \left(\frac{1}{4}\right)^2 [1 + (1 + \theta^\pi)^2 + (\theta^\pi)^2] \sigma_{\epsilon^\pi}^2\tag{13a}$$

$$\begin{aligned}E(\pi_{t,1} - \pi_{t,2})(\pi_{t,2} - \pi_{t,3}) &= \left(\frac{1}{4}\right)^2 E n_{t,q3}^\pi n_{t,q2}^\pi \\ &= \left(\frac{1}{4}\right)^2 E(\epsilon_t^\pi + \theta^\pi \epsilon_{t-1}^\pi)(\epsilon_{t-1}^\pi + \theta^\pi \epsilon_{t-2}^\pi) \\ &= \left(\frac{1}{4}\right)^2 \theta^\pi \sigma_{\epsilon^\pi}^2\end{aligned}\tag{13b}$$

and I use the same technique for estimating $\{\theta^y, \sigma_{\epsilon^y}^2\}$ in equation (10) using WEO's semi-annual releases on real GDP growth.

Once $\{\theta^\pi, \sigma_{\epsilon^\pi}^2\}$ and $\{\theta^y, \sigma_{\epsilon^y}^2\}$ are estimated, $\sigma_{n\pi}$ and σ_{ny} can be calculated from:

$$\begin{aligned}\sigma_{n\pi}^2 &= \sigma_{\epsilon^\pi}^2 (1 + (\theta^\pi)^2) \\ \sigma_{ny}^2 &= \sigma_{\epsilon^y}^2 (1 + (\theta^y)^2)\end{aligned}\tag{14}$$

and figure 2 (a) describes the cross-country variation in $\sigma_{n\pi}$ and σ_{ny} .

⁹One may argue that the second quarter inflation rate ($\pi_{t,q2}^n$) in the real-time data ($\pi_{t,0}$) should also be a forecast rather than a noisy observation since the WEO data are released at the beginning of the second and fourth quarters. This issue can be resolved by adding one more revision ($\pi_{t,4}$) to the list (equation (11)) and modifying the assumption as the following: it takes about a quarter for the IMF research department to receive and organize new data before releasing them on WEO database.

2.3 Robust pattern

Figure 2 (b) reports estimated $\tilde{\rho}$ for the sample countries with regard to the level of data uncertainty in each country, where $\tilde{\rho}$ is the coefficient on lagged interest rate in the Taylor rule equation (6). Note that $\tilde{\rho}$ is positively related with the level of data uncertainty. The inclusion of the lagged interest rate in the Taylor rule estimation can be rationalized by the substantial increase in adjusted- R^2 in figure 2 (c).

This cross-country pattern is robust to controlling currency pegs, country income levels, exchange rates, and federal funds rates, as reported in figure 3, 4 and 5.

Figure 3 shows that the cross-country pattern is robust in the subset of sample countries without hard currency peg. Panel (a) reports the scatter plot of the subset of the countries without a hard currency peg. Panel (b) reports the scatter plot of the subset of the countries without any currency peg (hard peg and soft peg). Panel (c) reports the scatter plot of the subset of the countries only with hard currency peg, and the scatter plot implies that the relationship does not hold among those countries with a hard currency peg. This is not surprising because domestic data uncertainty would not affect a central bank's monetary policy if its currency is perfectly pegged to another currency.

Figure 4 shows that the cross-country pattern is robust when country income levels are considered. Panel (a), (b) and (c) report the scatter plot within each subset of the sample countries with high, middle, and low income, respectively. The cross-country pattern among low-income countries is also consistent with the others if countries with a hard currency peg are eliminated.

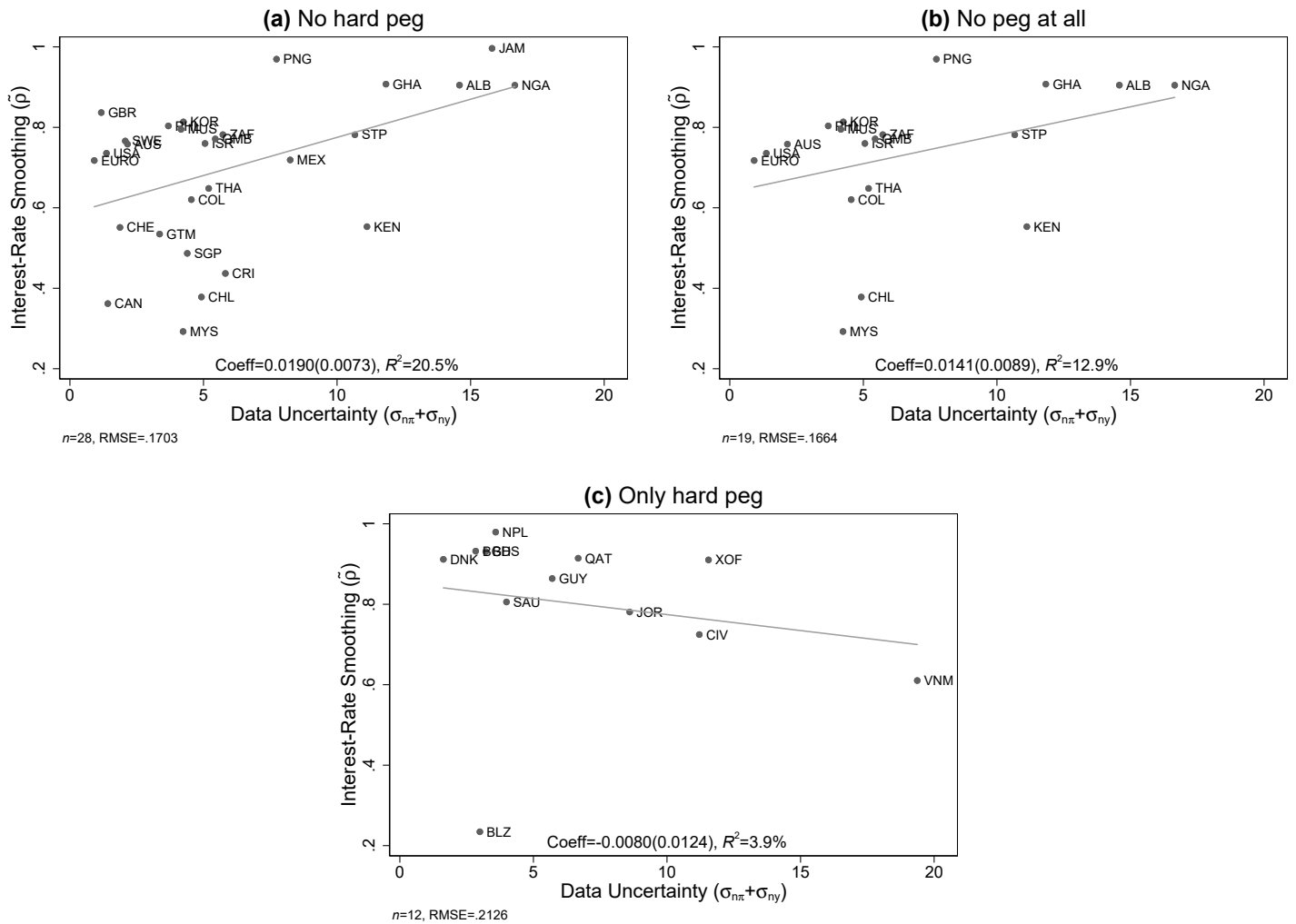
The cross-country pattern is robust when real effective exchange rates and federal funds rates are controlled. Real effective exchange rate and federal funds rate are added in addition to the baseline equation (equation (6)). I estimate the following specifications of the partial-adjustment Taylor rule for each country c using the least squares:

$$i_{t,c} = (1 - \tilde{\rho}_c)[k_c + \widetilde{g_{\pi,c}}\pi_{t+1,c} + \widetilde{g_{y,c}}y_{t,c} + \widetilde{g_{RER,c}}RER_t] + \tilde{\rho}_c i_{t-1,c} + \varepsilon_{t,c} \quad (15)$$

$$i_{t,c} = (1 - \tilde{\rho}_c)[k_c + \widetilde{g_{\pi,c}}\pi_{t+1,c} + \widetilde{g_{y,c}}y_{t,c} + \widetilde{g_{FFR,c}}FFR_t] + \tilde{\rho}_c i_{t-1,c} + \varepsilon_{t,c} \quad (16)$$

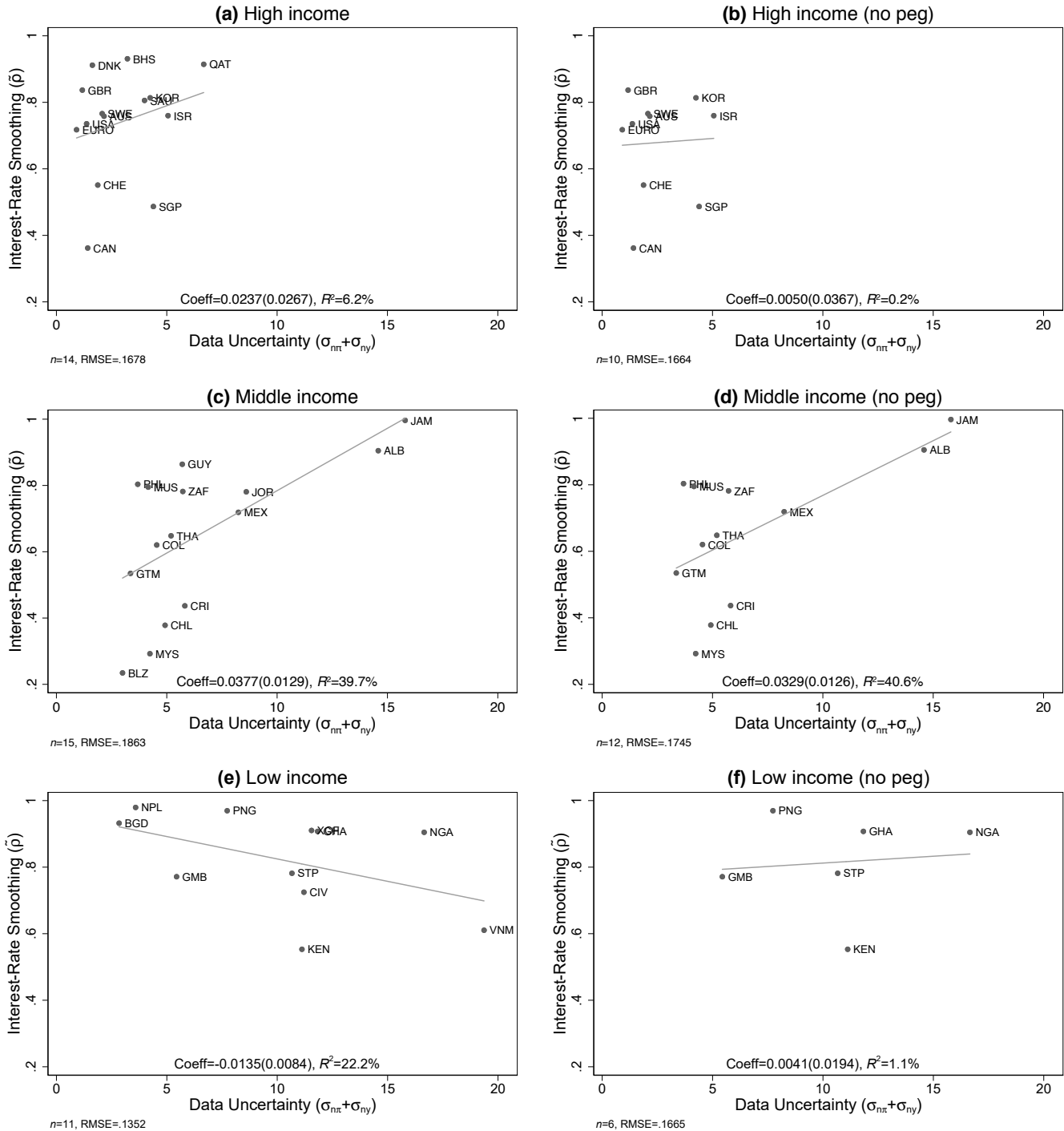
where $i_{t,c}$ is annualized monetary policy-related interest rate, $\pi_{t+n,c|t}$ is n -period ahead annualized CPI inflation rate forecast, $y_{t,c|t}$ is output gap measured by the percentage difference between real GDP and estimated potential GDP, $RER_{t,c}$ is real effective exchange rate (CPI base), and FFR_t is the annualized federal funds rate. The horizon of the infla-

Figure 3: Cross-country comparison of interest-rate smoothing ($\tilde{\rho}$) regarding currency pegs



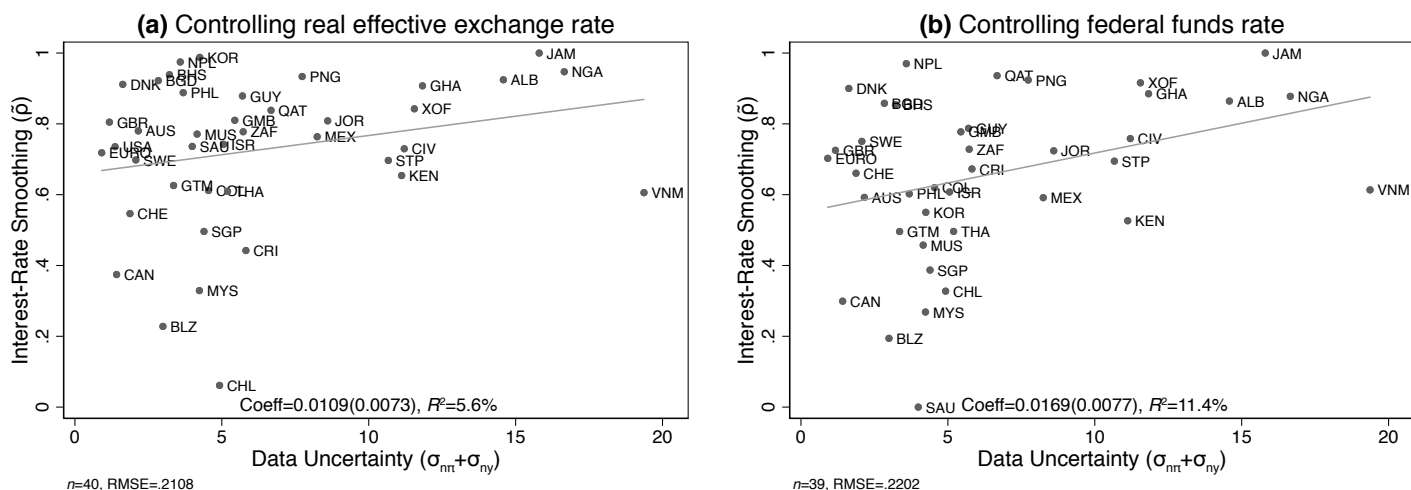
Note: These figures present subsets of the scatterplot in Figure 2 (b) with regard to exchange rate regimes. The classifications of exchange rate regimes are based on Shambaugh (2004) and Klein and Shambaugh (2008).

Figure 4: Cross-country comparison of interest-rate smoothing ($\tilde{\rho}$) by income level



Note: These figures present subsets of the scatterplot in Figure 2 (b) by country income level. The classifications of country income levels are based on the World Development Indicators database of the World Bank. Income is measured using gross national income (GNI) per capita, in U.S. dollars, converted from local currency using the World Bank Atlas method. The classifications of exchange rate regimes are based on Shambaugh (2004) and Klein and Shambaugh (2008).

Figure 5: Cross-country comparison of interest-rate smoothing ($\tilde{\rho}$) controlling real effective exchange rate and federal funds rate



Note: Panel (a) and (b) report estimated $\tilde{\rho}$ for 40 sample countries, based on the equation (15) and (16), respectively. Real effective exchange rate (RER_t) data are CPI-based and drawn from IMF’s IFS dataset, and the annualized federal funds rates (FFR_t) are drawn from FRED dataset by the Federal Reserve Bank of St. Louis. Panel (b) does not include the United States.

tion forecast is chosen to be one year. The scatter plots are reported in figure 5, and the cross-country pattern is robust in both panels.

3 Theoretical Framework

The robust cross-country pattern—the higher data uncertainty, the higher weight on lagged interest rate in the estimated Taylor rule—is further explored within a simple New Keynesian framework. I introduce a model that describes the economy in which a central bank minimizes its loss by choosing the monetary policy rule parameters. Three types of monetary policies are considered: monetary policy under perfect information, naive monetary policy, and monetary policy with learning.

3.1 Model setup

The model is taken from [Rudebusch and Svensson \(1999\)](#) with some modifications. The optimal policy rules are derived in a simple model of output and inflation:

$$\pi_t = \alpha_0 + \alpha_{\pi 1}\pi_{t-1} + \alpha_{\pi 2}\pi_{t-2} + \alpha_{\pi 3}\pi_{t-3} + \alpha_{\pi 4}\pi_{t-4} + \alpha_y y_{t-1} + \varepsilon_t \quad (17)$$

$$y_t = \beta_0 + \beta_{y1}y_{t-1} + \beta_{y2}y_{t-2} + \beta_r(\bar{i}_{t-1} - \bar{\pi}_{t-1}) + \eta_t \quad (18)$$

where π_t is inflation rate, y_t is output gap, i_t is interest rate, $\bar{i}_t = \frac{1}{4}\sum_{j=0}^3 i_{t-j}$, and $\bar{\pi}_t = \frac{1}{4}\sum_{j=0}^3 \pi_{t-j}$.¹⁰ The central bank is assumed to minimize the variation in inflation around its target π^* , in the output gap, and in changes in the interest rate. Expected loss equals the weighted sum of unconditional variances,

$$E[L_t] = Var[\bar{\pi}_t - \pi^*] + \lambda_y Var[y_t] + \lambda_i Var[\Delta i_t] \quad (19)$$

where $\Delta i_t = i_t - i_{t-1}$, and the parameters $\lambda_y \geq 0$ and $\lambda_i \geq 0$ are the relative weights on output stabilization and interest-rate smoothing, respectively, with respect to inflation stabilization. The central bank has the forward-looking policy rule in the form of:

$$i_t = (1 - \rho)(k + g_\pi \pi_{t+1|t} + g_y y_{t|t}) + \rho i_{t-1}, \quad (20)$$

where $\pi_{t+1|t}$ and $y_{t|t}$ are the central bank's forecast and inference on the future inflation, π_{t+1} , and the true output gap, y_t , and chooses the optimal values for g_π , g_y , and ρ . The Taylor rule sets the interest rate in quarter t on the basis of real-time noisy indicators on inflation, π_t^n , and output gap, y_t^n , which are defined as:

$$\pi_t^n = \pi_t + n_t^\pi \quad (21)$$

$$y_t^n = y_t + n_t^y \quad (22)$$

n_t^π and n_t^y are the contemporaneous measurement errors that plague the policymaker in real time, with standard errors $\sigma_{n\pi}$ and σ_{ny} , respectively, and uncorrelated with π_t and y_t . The measurement errors, n_t^π and n_t^y , are modeled as first-order moving average (MA(1)) processes:

$$n_t^\pi = \epsilon_t^\pi + \theta^\pi \epsilon_{t-1}^\pi \quad (23)$$

¹⁰The model is in quarterly frequency.

$$n_t^y = \epsilon_t^y + \theta^y \epsilon_{t-1}^y \quad (24)$$

where ϵ_t^π and ϵ_t^y are normally distributed with mean zero and variance $\sigma_{\epsilon^\pi}^2$ and $\sigma_{\epsilon^y}^2$, respectively.

Depending on the central bank's information and policy type, I introduce three monetary policy cases: perfect information and naive monetary policy as benchmark cases and monetary policy with learning for the main analysis. For each monetary policy case, the expectation on future inflation rate ($\pi_{t+1|t}$) and the expectation on current GDP gap ($y_{t|t}$) differ, and these lead to different reactions of the central bank.

3.2 Perfect information

Under the assumption that the central bank always has perfect information (i.e., $\pi_{t|t} = \pi_t$ and $y_{t|t} = y_t$ in equation (20)), optimal policy is described by:

$$i_t = (1 - \rho^P)(k + g_\pi^P E_t[\pi_{t+1}] + g_y^P y_t) + \rho^P i_{t-1} \quad (25)$$

where $E_t[\pi_{t+1}]$ is the rational expectation on π_{t+1} using all current and past perfect information at time t . The central bank optimally chooses its response parameters, g_π^P , g_y^P and ρ^P , to minimize the loss function (equation (19)). These response parameters are invariant to the magnitude of noise in the data because the central bank always observes and responds to the true data.

3.3 Naive monetary policy

The policymaker may take the noisy indicators at face value without making any inference about the true inflation rate and GDP gap (i.e., $\pi_{t|t} = \pi_t^n$ and $y_{t|t} = y_t^n$ in equation (20)). The optimal policy is described by:

$$i_t = (1 - \rho^N)(k + g_\pi^N E_t[\pi_{t+1}] + g_y^N y_t^n) + \rho^N i_{t-1} \quad (26)$$

where $E_t[\pi_{t+1}]$ is the rational expectation on π_{t+1} using all available information at time t , which can be derived from the Phillips curve (equation (17)) using current and past noisy data.¹¹ The central bank optimally chooses rule parameters, g_π^N , g_y^N and ρ^N , such that

¹¹At the time of decision, the central bank's information set includes recent four quarters of noisy indicators $\{\pi_t^n, \pi_{t-1}^n, \pi_{t-2}^n, \pi_{t-3}^n; y_t^n, y_{t-1}^n, y_{t-2}^n, y_{t-3}^n\}$ and historical true data $\{\pi_{t-4}, \pi_{t-5}, \dots, \pi_{t-\infty};$

minimize the loss function (equation (19)). Since the true data are not observable, the central bank chooses different rule parameters as data become noisier.

3.4 Monetary policy with learning

Naive monetary control is efficient in the absence of noise but is inefficient when noise is present in the data since the policymaker can reduce her loss using forecastable components in the noise process. By implementing the Kalman filter, the central bank makes optimal inferences on the inflation rate and GDP gap given available information.

Kalman filtering with noisy information

The model described in section (3.1) has a state-space representation,

$$X_{t+1} = AX_t + Bi_t + \boldsymbol{\nu}_{t+1} \quad (27)$$

The 10×1 vector X_t of state variables, the 10×10 matrix A , the 10×1 column vector B , and the 10×1 column disturbance vector $\boldsymbol{\nu}_t$ are given by

$$X_t = \begin{bmatrix} 1 \\ \pi_t \\ \pi_{t-1} \\ \pi_{t-2} \\ \pi_{t-3} \\ y_t \\ y_{t-1} \\ i_{t-1} \\ i_{t-2} \\ i_{t-3} \end{bmatrix}, A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_0 & \alpha_{\pi 1} & \alpha_{\pi 2} & \alpha_{\pi 3} & \alpha_{\pi 4} & \alpha_y & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_0 & -\beta_r/4 & -\beta_r/4 & -\beta_r/4 & -\beta_r/4 & \beta_{y1} & \beta_{y2} & \beta_r/4 & \beta_r/4 & \beta_r/4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \beta_r/4 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\nu}_t = \begin{bmatrix} 0 \\ \varepsilon_t \\ 0 \\ 0 \\ 0 \\ \eta_t \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The central bank's observation equation is given as

$$Z_t = CX_t + \boldsymbol{w}_t \quad (28)$$

where

$$Z_t = \begin{bmatrix} \pi_t^n \\ y_t^n \end{bmatrix}, C = \begin{bmatrix} e_2 \\ e_6 \end{bmatrix}, \boldsymbol{w}_t = \begin{bmatrix} n_t^\pi \\ n_t^y \end{bmatrix},$$

$y_{t-4}, y_{t-5}, \dots, y_{t-\infty}$ assuming that the noisy indicators are revised and become true values after four quarters.

and e_j denotes a 1×10 row vector with element j equal to unity and all other elements equal to zero. Z_t is the observation vector which consists of noisy indicators, C_t is the observation model which maps the true state space into the observed space, and w_t is the vector of observation noises.¹²

Optimal Kalman gain K (10×2 matrix) and predicted error covariance $P_{t|t-1}$ (10×10 matrix) are specified as

$$\begin{aligned} K &= P_{t|t-1}C^T(CP_{t|t-1}C^T + V_w)^{-1} \\ P_{t|t-1} &= A(P_{t|t-1} - KCP_{t|t-1})A^T + V_\nu \end{aligned} \quad (29)$$

where V_ν (10×10 matrix) and V_w (2×2 matrix) are variance-covariance matrices of ν_t and w_t , respectively.

The central bank's optimal inference $X_{t|t}$ and forecast $X_{t+1|t}$ are in recursive form as

$$\begin{aligned} X_{t|t} &= X_{t|t-1} + K(Z_t - Z_{t|t-1}) \\ &= (I - KC)AX_{t-1|t-1} + (I - KC)Bi_{t-1} + KZ_t \end{aligned} \quad (30)$$

$$X_{t+1|t} = AX_{t|t} + Bi_t \quad (31)$$

where $X_{t|t-1}$ denotes predicted (a priori) state estimate and $Z_t - Z_{t|t-1}$ denotes innovation (measurement pre-fit residual).¹³ In each period, by observing noisy data, the central bank learn about the noise process and decide how much weight (K) will be placed on upcoming noisy observation while the rest of the weight ($I - K$) will be placed on the prior inference.

Optimal monetary policy with learning

The central bank's policy rule in equation (20) can be written as

$$i_t = (1 - \rho)(k + GX_{t|t}) + \rho i_{t-1} \quad (32)$$

where

$$G = \begin{bmatrix} g_\pi & g_y \end{bmatrix} \begin{bmatrix} e_2 A \\ e_6 \end{bmatrix} \quad (33)$$

¹²Standard Kalman filter assumes zero-mean Gaussian white noise for w_t . In this paper, I introduce an auxiliary random process so that w_t follows MA(1) within the Kalman filter. Please refer to appendix (A) for more details.

¹³Note that $Z_{t|t-1} = CX_{t|t-1}$ and $X_{t|t-1} = AX_{t-1|t-1} + Bi_{t-1}$.

and e_j denotes a 1×10 row vector with element j equal to unity and all other elements equal to zero.¹⁴

The central bank optimally chooses response coefficients, g_π , g_y and ρ , minimizing the loss function (equation (19)) based on the central bank's inferences ($X_{t|t}$).

4 Results

Using the model described above, I estimate the optimal monetary policy rule parameters, g_π , g_y and ρ , under the three information and policy cases. Then, I run reduced-form Taylor rule regressions on simulated data, which are generated by the same model, using the specification in section (2.1). I show that the optimal level of gradual adjustment (ρ) derived from the model and the reduced-form estimate on the lagged interest rate ($\tilde{\rho}$) move toward the opposite directions when the central bank faces more data uncertainty. I describe the results and explain the underlying intuition in this section.

The model in the previous section is calibrated using the following parameters. The parameters in Phillips and IS curves (equations (23) and (24)) are estimated in the data, and the parameters in the loss function are assumed to be $\lambda_y = 1$ and $\lambda_i = 0.5$, as appear in Rudebusch and Svensson (1999).¹⁵ Later in this section, I consider the loss function with $\lambda_i = 0$ in a purpose of muting the central bank's cautious motive. The parameters in MA(1) noise processes (equations (23) and (24)) are estimated for 40 sample countries in the historical data, and, for example, they are $\theta^\pi = 0.51$ and $\theta^y = 0.58$ in the United States.

4.1 Optimal monetary policy under data uncertainty

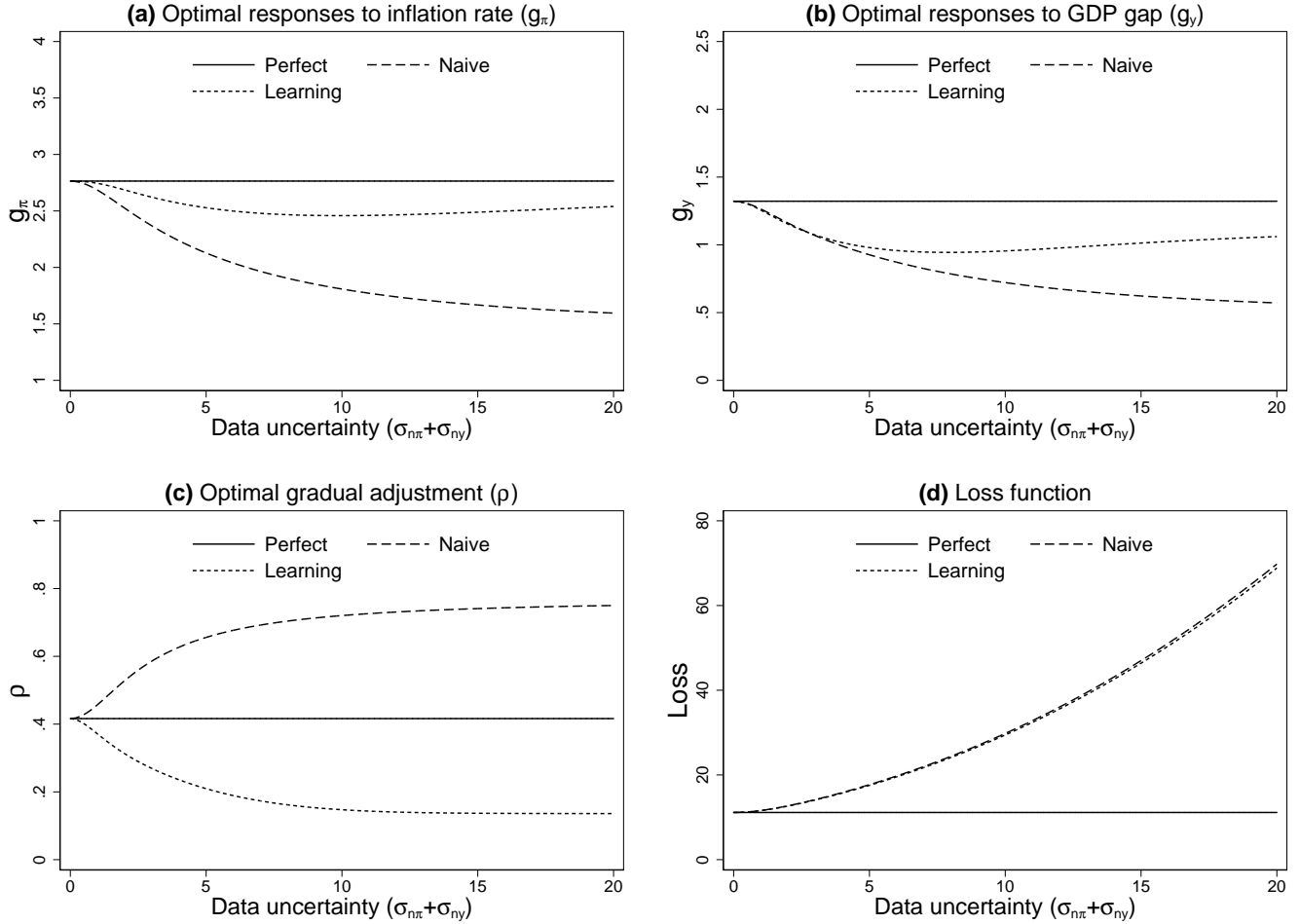
Depending on information and policy type, the central bank optimally responses to noise in the data. The optimal rule parameters, g_π , g_y and ρ , are estimated using the grid search and reported in figure 6. The optimal responses and the associated loss under perfect information, naive policy, and learning policy are described as solid lines, long-dashed lines, and short-dashed lines, respectively.

Under the perfect information case, the central bank observes data without error and responds to the true inflation rate and GDP gap. The optimal responses and loss are inde-

¹⁴Note that $\pi_{t+1|t} = e_2 X_{t+1|t} = e_2 (A X_{t|t} + B i_t) = e_2 A X_{t|t}$ and $y_{t|t} = e_6 X_{t|t}$.

¹⁵It is called a "strict" inflation targeting if only inflation enters the loss function ($\lambda_y = \lambda_i = 0$). A "flexible" inflation targeting allows other goal variables in the loss function. See Rudebusch and Svensson (1999) for details.

Figure 6: Optimal rule parameters and loss function subject to data uncertainty in US



Note: Figure 6 reports the responses of optimal monetary policy rule parameters, g_{π} , g_y and ρ , under perfect information, naive monetary policy, and monetary policy with learning, to noise in the data. The parameters are estimated with grid search.

The estimated parameters in Phillips and IS curves (equation (17) and (18)) are described as:

$$\pi_t = \underset{(0.09)}{0.08} + \underset{(0.08)}{0.67}\pi_{t-1} - \underset{(0.10)}{0.08}\pi_{t-2} + \underset{(0.10)}{0.29}\pi_{t-3} + \underset{(0.08)}{0.12}\pi_{t-4} + \underset{(0.03)}{0.15}y_{t-1} + \varepsilon_t, \sigma_{\varepsilon\pi} = 1.007$$

$$y_t = \underset{(0.10)}{0.19} + \underset{(0.08)}{1.17}y_{t-1} - \underset{(0.08)}{0.27}y_{t-2} - \underset{(0.03)}{0.09}(\bar{l}_{t-1} - \bar{\pi}_{t-1}) + \eta_t, \sigma_{\varepsilon y} = 0.822$$

and the estimated parameters in MA(1) noise processes (equations (23) and (24)) are $\theta^{\pi} = 0.51$ and $\theta^y = 0.58$.

pendent of data uncertainty. On the other hand, the central bank's responses, g_π and g_y , under naive and learning policies generally decrease in data uncertainty. Panel (d) reports that loss strictly increases in data uncertainty under both policy cases.

The optimal degree of gradual adjustment (ρ), which is the parameter of interest, increases in data uncertainty under the naive policy and decreases in data uncertainty under the learning policy, as described in panel (c). The optimal degree of gradual adjustment (ρ) is determined by the tradeoff in the loss function, which consists of three components: $Var[\bar{\pi}_t - \pi^*]$, $Var[y_t]$, and $Var[\Delta i_t]$. Each component varies over ρ , and the central bank orchestrates the contribution of each component by choosing the optimal ρ that minimize the loss function. Note that $Var[\bar{\pi}_t - \pi^*]$ and $Var[y_t]$ increase in ρ , while $Var[\Delta i_t]$ decreases in ρ .¹⁶ If $\lambda_i = 0$, which implies that the loss is independent of interest-rate volatility, then the optimal ρ is zero regardless of data uncertainty in all three cases, because both $Var[\bar{\pi}_t - \pi^*]$ and $Var[y_t]$ uniformly increase in ρ . If $\lambda_i > 0$, then the optimal ρ is greater than or equal to zero, and it increases in data uncertainty under the naive policy and decreases in data uncertainty under the learning policy.

The main reason why ρ moves differently under different policy case can be explained with the relative contribution of $Var[\Delta i_t]$ to the loss: the relative contribution of $Var[\Delta i_t]$ in the loss function increases in data uncertainty under the naive policy and decreases in data uncertainty under the learning policy, as described in figure 7. Using the Taylor rule in equation (32), $Var[\Delta i_t]$ can be written as

$$\begin{aligned} Var[\Delta i_t] &= Var \left[[(1 - \rho)(k + GX_{t|t}) + \rho i_{t-1}] - [(1 - \rho)(k + GX_{t-1|t-1}) + \rho i_{t-2}] \right] \\ &= Var \left[(1 - \rho)(G\Delta X_{t|t}) + \rho \Delta i_{t-1} \right] \end{aligned} \quad (34)$$

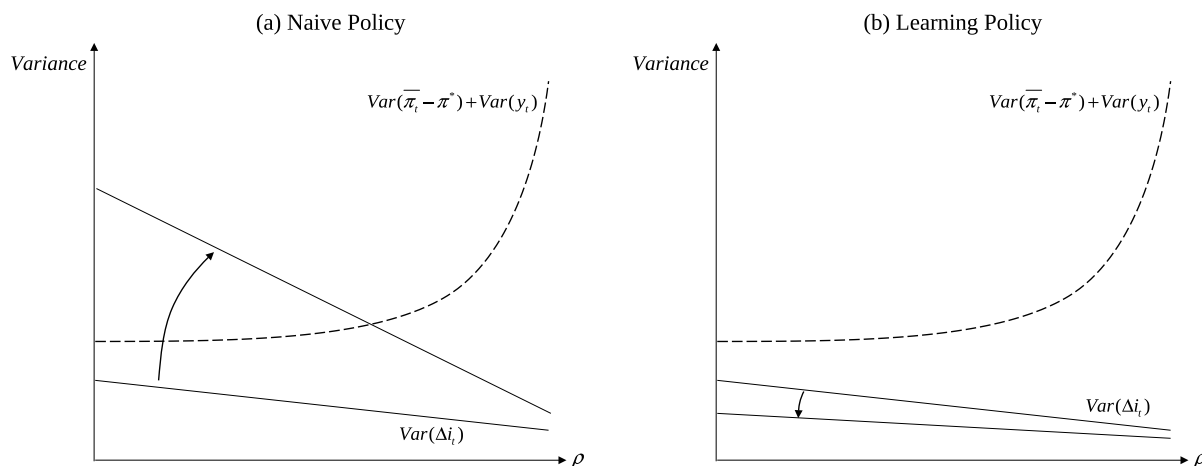
where $\Delta X_{t|t} = X_{t|t} - X_{t-1|t-1}$ and $\Delta i_{t-1} = i_{t-1} - i_{t-2}$. Since $Var[\Delta i_t] = Var[\Delta i_{t-1}]$ in steady state,

$$Var[\Delta i_t] = \frac{1 - \rho}{1 + \rho} Var(G\Delta X_{t|t}) + \frac{2\rho}{1 + \rho} Cov(G\Delta X_{t|t}, \Delta i_{t-1}) \quad (35)$$

where $\frac{1 - \rho}{1 + \rho}$ is a decreasing function and $\frac{2\rho}{1 + \rho}$ is an increasing function in ρ . Fixing everything else constant, $Var[\Delta i_t]$ decreases in ρ because $Var(G\Delta X_{t|t})$ is positive and

¹⁶If ρ approaches to one, then the interest rate process becomes close to a unit root process, whose variance diverges to infinity. Noting that π_t and y_t are functions of current and lagged interest rates in equation (17) and (18), $Var[\bar{\pi}_t - \pi^*]$ and $Var[y_t]$ increase in ρ . On the other hand, $Var[\Delta i_t]$ decreases in ρ since high ρ reduces volatility in interest rates.

Figure 7: Decomposition of the loss function



Note: These diagrams describe the shifts of $Var(\Delta i_t)$ subject to data uncertainty under naive policy and learning policy, respectively, adjusting the scales. Higher data uncertainty increases the slope of $Var(\Delta i_t)$ curve under the naive policy and decreases the slope of $Var(\Delta i_t)$ under the learning policy. The shape of $[Var(\bar{\pi}_t - \pi^*) + Var(y_t)]$ curves relatively does not change much under both policies.

$Cov(G\Delta X_{t|t}, \Delta i_{t-1})$ is negative in the absence of noise.¹⁷ When the data become noisier, $Var(G\Delta X_{t|t})$ and $Cov(G\Delta X_{t|t}, \Delta i_{t-1})$ change differently under different policy.

Naive monetary policy

The optimal choice of ρ increases in data uncertainty because $Var[\Delta i_t]$ curve in the figure 7 becomes steeper due to the following reason. When there is data uncertainty, the central bank's responses, g_π and g_y , become less useful since they are the responses not only to the true inflation and GDP gap but also to the noise. If the central bank keeps ρ fixed, it causes high fluctuations in interest rates resulting in greater $Var[\Delta i_t]$ in the loss function and a steeper $Var[\Delta i_t]$ curve in the figure 7 (a). Therefore, the central bank raises ρ to stabilize the interest rates.

An algebraic explanation is the following. Note that $X_{t|t} = Z_t$ under the naive policy, then equation (35) becomes

$$Var[\Delta i_t] = \frac{1-\rho}{1+\rho} Var(G\Delta Z_t) + \frac{2\rho}{1+\rho} Cov(G\Delta Z_t, \Delta i_{t-1}) \quad (36)$$

¹⁷For $0 \leq \rho \leq 1$, $\frac{1-\rho}{1+\rho}$ is convex and monotone decreasing in ρ from 1 to 0, and $\frac{2\rho}{1+\rho}$ is concave and monotone increasing in ρ from 0 to 1.

$Var[\Delta i_t]$ curve in the figure 7 (a) becomes steeper as data uncertainty increases because: (i) $\frac{1-\rho}{1+\rho}$ decreases in ρ and $Var(G\Delta Z_t)$ increases very much in noise; (ii) the second term, $\frac{2\rho}{1+\rho}$ increases in ρ and $Cov(G\Delta Z_t, \Delta i_{t-1})$ decreases in noise (from negative to further negative), when evaluated at empirically relevant parameter values. The marginal benefit of increasing ρ becomes large.

Monetary policy with learning

The optimal responses to the inflation rate and GDP gap are moderate under learning policy, compared to those under the naive policy, because the central bank's inferences on the true data are more informative than noisy data. The loss is also smaller under the learning policy. One interesting finding is that the optimal interest-rate smoothing parameter, ρ , decreases in data uncertainty.

The central bank's optimal choice of ρ decreases in data uncertainty because the learning process (i) can serve to effectively filter the noise, and (ii) tease out from the persistence of the noise additional information about the true data. The learning process helps the central bank to make decisions based on the much less noisy information. The second point needs to be more discussed. The noise in the data is assumed to follow MA(1) process, and the estimated θ^π and θ^y are both positive in almost all countries in the sample, implying a positive autocorrelation in the noise. A simple thought experiment helps in understanding how the central bank learns the additional information from the persistent noise.

Let's say there was a positive noise shock in the past period, then that means the optimal policy in that period resulted in a higher interest rate than it would be under perfect information. Due to a positive autocorrelation in the noise process, the chance of getting another positive noise shock in the current period is more likely. If the central bank increases or keeps ρ constant, then the interest rate in the current period will likely be higher than the central bank wants because the interest rate in the past period is already high. Therefore, the central bank reduces ρ .

The same logic applies to the other way. If there was a negative noise shock in the past period, then that means the optimal policy in that period resulted in a lower interest rate than it would be under perfect information. Due to a positive autocorrelation in the noise process, the chance of getting another negative noise shock in the current period is more likely. If the central bank increases or keeps ρ constant, then the interest rate in the current period will likely be lower than the central bank wants because the interest rate in the past period is already low. Therefore, the central bank reduces ρ . In both cases, the central bank

can reduce interest rate volatility by lowering ρ .

This thought experiment can be supported by figure 8 (a). If noise is persistent, or the noise process has a positive autocorrelation ($\theta > 0$), then optimal ρ decreases in data uncertainty. On the other hand, if noise is not persistent, or the noise process has a negative autocorrelation ($\theta < 0$), then optimal ρ increases in data uncertainty.¹⁸

As mentioned above, the optimal degree of gradual adjustment (ρ) is determined by the tradeoff in the loss function. The central bank's optimal choice of ρ decreases in data uncertainty because $Var[\Delta i_t]$ curve in the figure 7 (b) becomes flatter as data uncertainty increases. An algebraic explanation is the following.

Note that $X_{t|t}$ under the learning policy is given in equation (30), then equation (35) is

$$Var[\Delta i_t] = \frac{1 - \rho}{1 + \rho} Var(G\Delta X_{t|t}) + \frac{2\rho}{1 + \rho} Cov(G\Delta X_{t|t}, \Delta i_{t-1}) \quad (37)$$

where

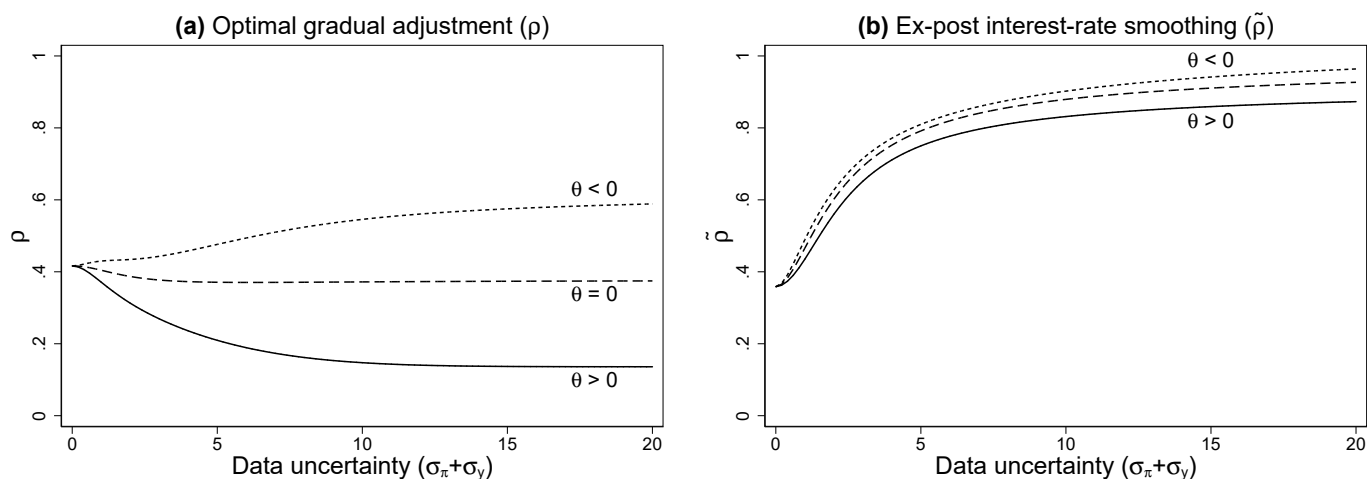
$$\Delta X_{t|t} = (I - KC)A\Delta X_{t-1|t-1} + (I - KC)B\Delta i_{t-1} + K\Delta Z_t \quad (38)$$

$Var[\Delta i_t]$ curve in the figure 7 (b) becomes flatter as noise increases because (i) $\frac{1-\rho}{1+\rho}$ decreases in ρ and $Var(G\Delta X_{t|t})$ increases little in noise; (ii) the second term, $\frac{2\rho}{1+\rho}$ increases in ρ and $Cov(G\Delta X_{t|t}, \Delta i_{t-1})$ increases very much in noise (from negative to positive), when evaluated at empirically relevant parameter values. The marginal benefit of increasing ρ becomes small.

Under the learning policy, $Var(G\Delta X_{t|t})$ increases little in noise because the learning process effectively filters the noise. The variance is not zero since, even if the learning process is effective, it cannot completely eliminate the noise in the data. $Cov(G\Delta X_{t|t}, \Delta i_{t-1})$ increases very much in noise because the learning process helps the central bank to gain some information from the persistence in the noise. If noise is not persistent, or the noise process has a negative autocorrelation ($\theta < 0$), then this covariance decreases in noise, same as in the naive policy case.

¹⁸When the noise is white noise ($\theta = 0$), the optimal ρ decreases when data uncertainty is small because the central bank is not sure if the shock is from the noise or from Phillips and IS curves. Note that Phillips and IS curves have shock terms, ε_t and η_t , and their standard errors are 1.007 and 0.822, respectively, in the United States. If data uncertainty is big enough, then the central bank believes that most of the shock is coming from data uncertainty, rather than Phillips and IS curves, and the optimal ρ does not change.

Figure 8: Optimal gradual adjustment (ρ) and ex-post interest-rate smoothing ($\tilde{\rho}$) with varying signs of θ under learning policy



Note: Panel (a) reports the estimated optimal gradual adjustment (ρ) under learning policy when θ is positive, zero, and negative. Panel (b) reports the reduced-form estimate of interest-rate smoothing ($\tilde{\rho}$) for each θ .

4.2 Ex-post estimates using simulated data

I run reduced-form Taylor rule regressions on simulated data from the model and show that, when the central bank faces more data uncertainty, the reduced-form estimate on the lagged interest rate ($\tilde{\rho}$) may move differently from the optimal level of gradual adjustment (ρ) derived from the model.

Using the optimal responses described in the previous section, I simulate the model and generate 100,000 observations (and the first 20,000 observations are dropped) given each level of data uncertainty from 0 to 20. Using the simulated data, I replicate the empirical finding that higher data uncertainty leads to more interest-rate smoothing (i.e., the greater $\tilde{\rho}$) by estimating the reduced-form specification in section (2.1), that is

$$\hat{i}_t = (1 - \tilde{\rho})[k + \tilde{g}_\pi E_t \pi_{t+1} + \tilde{g}_y y_t] + \tilde{\rho} \hat{i}_{t-1} + \varepsilon_t \quad (39)$$

where the expected inflation rate, $E_t \pi_{t+1}$, is measured by the four quarter average inflation rate in percent. The reduced-form estimates, \tilde{g}_π , \tilde{g}_y and $\tilde{\rho}$, are based on *ex post* observation, and they are reported in figure (9).

Note that the central bank's optimal inference $X_{t|t}$ is correlated with i_{t-1} . Combining

equation (30) and (32), the policy rule under the learning policy can be expressed as

$$i_t = (1 - \rho)(k + G[(I - KC)AX_{t-1|t-1} + (I - KC)Bi_{t-1} + KZ_t]) + \rho i_{t-1} \quad (40)$$

and the reduced-form estimate on i_{t-1} consists of not only ρ but also coefficients on the correlating components in the central bank's belief (let's call it δ , then $\tilde{\rho} = \rho + \delta$). Figure (9) reports responses of the reduced-form estimates, \tilde{g}_π , \tilde{g}_y and $\tilde{\rho}$, with regard to data uncertainty.

Under the learning policy, the optimal level of gradual adjustment (ρ), calibrated in the model, decreases in data uncertainty, as described in figure 6 (c). On the other hand, the reduced-form estimate on the lagged interest rate ($\tilde{\rho}$) increases in data uncertainty. This explains the *intrinsic interest-rate smoothing* in the literature: the higher the data uncertainty, the greater the reduced-form estimate on the lagged interest rate. Interest-rate smoothing is mostly a consequence of the *learning process*, rather than central banks' *cautious motive*. Central banks' real-time belief about true data is indistinguishable from the cautious motive in ex-post data, both of which are largely picked up by the Taylor rule's reduced-form estimate on the lagged interest rate.

4.3 No-cautious-motive constraint

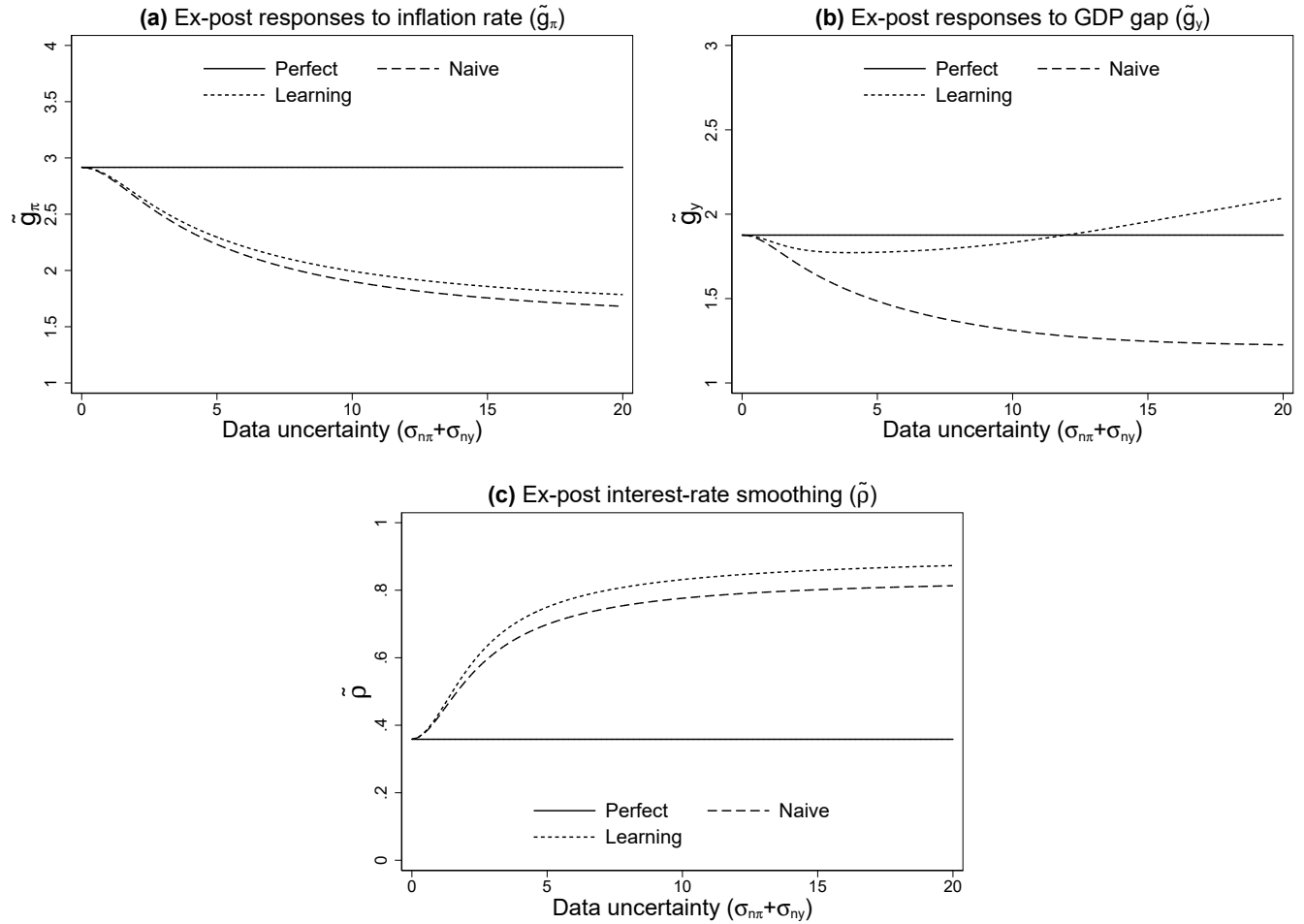
If the central bank has $\lambda_i = 0$, this implies no cautious motive ($\rho = 0$), and the monetary policy rule with learning in equation (32) becomes:

$$i_t = k + GX_{t|t} \quad (41)$$

Under the additional restriction, the optimal rule parameters, g_π and g_y , can be estimated from the model, and the reduced-form estimates, \tilde{g}_π , \tilde{g}_y and $\tilde{\rho}$ in equation (39), can be obtained from the simulated data. The estimation results are reported in figure 10. Panel (a) reports the optimal responses of the rule parameters under the constraint, and panel (b) reports the ex-post estimates of the rule parameters. Panel (c) reports the optimal level of caution (ρ) and interest-rate smoothing parameter ($\tilde{\rho}$). Panel (d) reports the loss under each monetary policy given the additional constraint.

Since there is no cautious motive ($\lambda_i = 0$), ρ becomes zero, and interest-rate smoothing ($\tilde{\rho}$) is completely induced by the learning process ($\tilde{\rho} = \delta$). The reduced-form estimate on the lagged interest rate ($\tilde{\rho}$) increases in data uncertainty even though the central bank does

Figure 9: Reduced-form estimates on ex-post rule parameters using simulated data



Note: Figure 9 reports reduced-form estimates of the Taylor rule (equation (39)) based on simulated data from the model. Estimated \tilde{g}_π , \tilde{g}_y and $\tilde{\rho}$ are reported under perfect information, naive monetary policy, and monetary policy with learning.

not conduct gradual adjustment (i.e., $\rho = 0$). A positive estimate on lagged interest rate does not necessarily mean that the central bank gradually adjusts the interest rates.

I show interest-rate smoothing ($\tilde{\rho}$) to exist and increase in the level of data uncertainty even if central banks' cautious motive is muted by excluding the variance of changes in interest rates from the loss function. Interest-rate smoothing ($\tilde{\rho}$) can be fully endogenized in the model by the learning process (δ) when there is no cautious motive. If the central bank has some caution, then interest-rate smoothing ($\tilde{\rho}$) can be decomposed into the learning process (δ) and gradual adjustment (ρ).

4.4 Cross-country analysis

Country-specific data uncertainty parameters $\{\sigma_{n\pi}, \sigma_{ny}, \theta^\pi, \theta^y\}$ are estimated from the historical WEO data. Each country has different level and relative composition of $\sigma_{n\pi}$ and σ_{ny} , and the levels of noise persistence, θ^π and θ^y , also differ across countries. Given data uncertainty in each country, the model is calibrated to estimate the optimal level of gradual adjustment parameter (ρ) and its *ex-post* reduced-form estimate ($\tilde{\rho}$) under the naive and learning policy, and these are reported in figure 11.

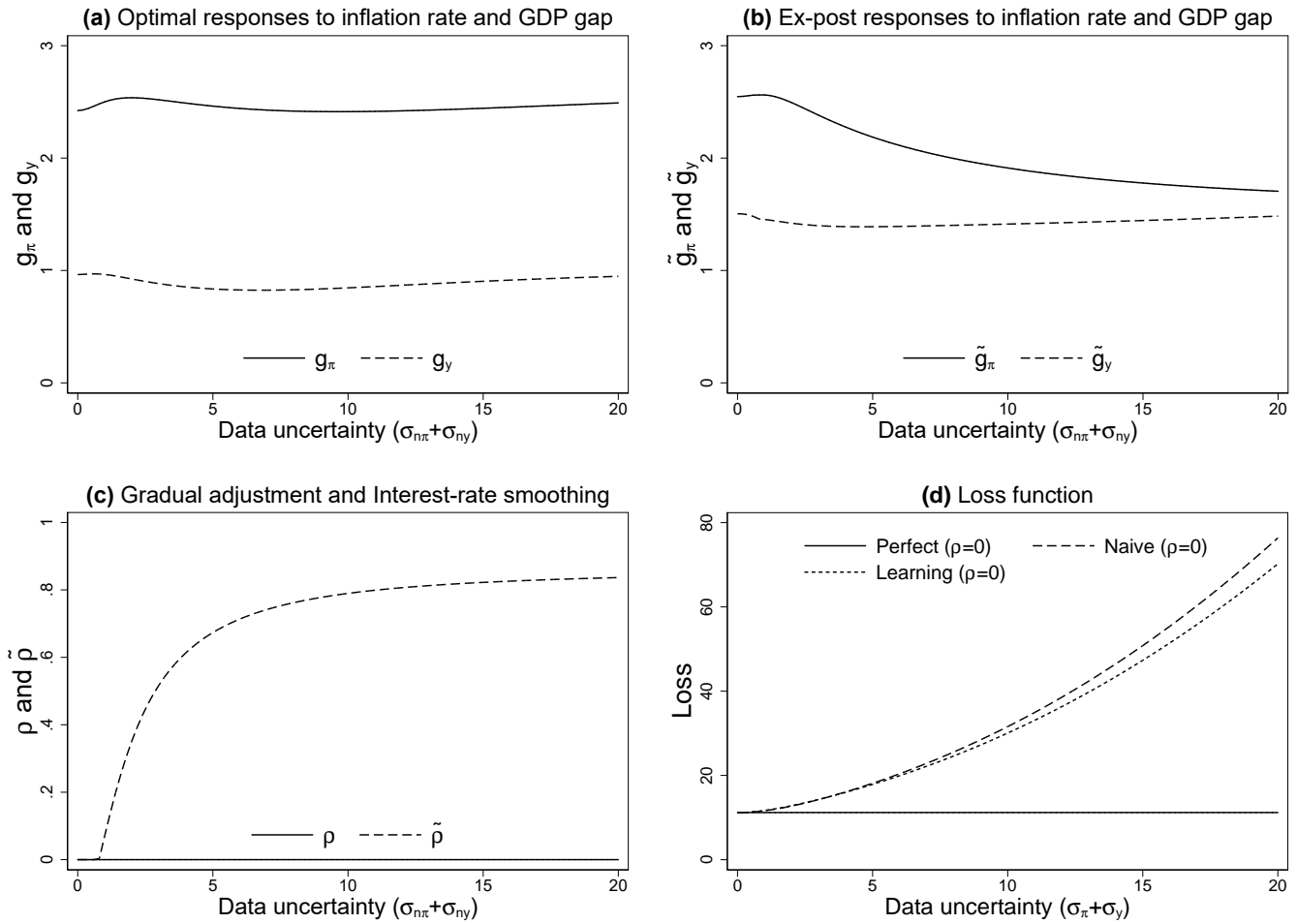
Panel (a) and (b) show that, under the naive policy, both ρ and $\tilde{\rho}$ increase in data uncertainty. On the other hand, panel (c) and (d) show that, under the learning policy, ρ and $\tilde{\rho}$ move differently — ρ decreases in data uncertainty and $\tilde{\rho}$ increases in data uncertainty. Given that the learning model described in this paper is relatively simple compared to the efforts and investment in central banks, panel (d) is most likely the cross-country pattern that we observe in figure 2 (b).¹⁹

4.5 Policy implications

We can think about how central banks can improve their monetary policy under data uncertainty. There are two approaches we can think of: improving learning ability and improving data quality. This research provides a framework that helps analyzing these competing options. We consider two benchmark cases in addition to the learning policy. First, central bank always observes perfect information. Second, central bank observes noisy data and naively take the face value without making any inference. We compare the welfare

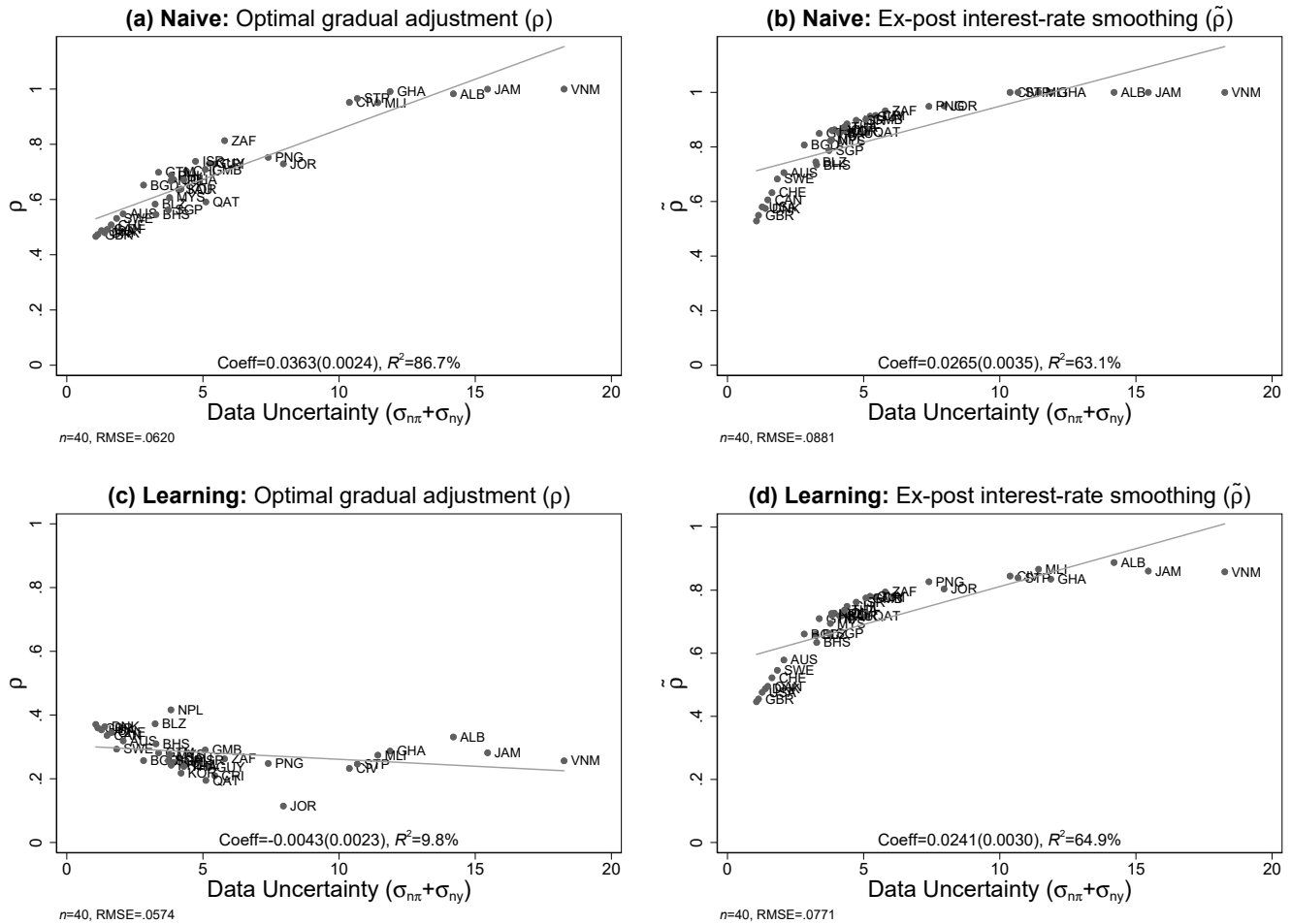
¹⁹The limitation of this practice is that I use the same model to estimate ρ and $\tilde{\rho}$ for all sample countries. Estimating the Phillips and IS curves for each country requires further discussions, and it is left for future study.

Figure 10: Optimal responses and ex-post estimates under learning muted cautious motive



Note: Figure 10 reports the optimal responses and ex-post estimates of the Taylor rule when central bank's cautious motive is muted ($\lambda_i = 0$). The loss function does not include the variance of the changes in interest rates, and this leads ρ become zero. Panel (a) reports optimal monetary policy rule parameters, g_π and g_y , under monetary policy with learning. The parameters are estimated with grid search. Panel (b) reports reduced-form estimates of the Taylor rule (equation (39)) based on simulated data from the model. Estimated \tilde{g}_π and \tilde{g}_y are reported under monetary policy with learning. Panel (c) reports ρ and $\tilde{\rho}$ under monetary policy with learning.

Figure 11: Cross-country estimates and optimal parameters under naive and learning policy



Note: Country-specific data uncertainty parameters $\{\sigma_{n\pi}, \sigma_{ny}, \theta^\pi, \theta^y\}$ are estimated from the historical WEO data, and the same model is calibrated for each country. Country-specific level of persistence in the noise and ratio in data uncertainty are allowed.

loss under the learning policy and the naive policy at different levels of data uncertainty and demonstrate that learning policy is always better than the naive policy.²⁰ The gain from learning is increasing and convex in data uncertainty. Given the cost information of each option, we can conduct a cost-benefit analysis and decide which option is more cost-effective.

5 Conclusion

The conventional belief about the relationship between the level of data uncertainty and interest-rate smoothing has been re-examined. The cross-country comparison of coefficients on the lagged interest rate supports the conventional belief, and I could replicate this with an appropriate calibration of the [Rudebusch and Svensson \(1999\)](#) model. However, the interest-rate smoothing is present in the reduced-form estimation of the Taylor rule, not because central banks gradually adjust the interest rates nor central banks are more cautious in the face of data uncertainty, but because the reduced-form estimates are obtained with *ex-post* data, in which the central banks' beliefs are not distinguishable, and because the central banks' inference about the true data is correlated with past interest rates.

This paper distinguishes between interest-rate smoothing and central banks' gradual interest-rate adjustment and demonstrates that the two may not move together. Whereas the conventional view in the literature considers interest-rate smoothing as evidence of central banks' cautious and gradual adjustment, I show that central banks' gradual interest-rate adjustment is directly induced by central banks' preference to avoid interest-rate surprises, that is, the variance of the changes in interest rates in the loss function.

This paper endogenizes interest-rate smoothing as a result of the central banks' learning process. I show that interest-rate smoothing exists and increases with data uncertainty, even if central banks' cautious motive is completely muted, by taking out the variance of the changes in interest rates from the loss function. Much of interest-rate smoothing comes from the learning process, and it can be fully endogenized by the cautious motive and the learning process in the model.

²⁰See loss functions in figure 6 (d) and figure 10 (d).

References

- Bernanke, Ben S.** 2004. “Gradualism.” Remarks at an economics luncheon co-sponsored by the Federal Reserve Bank of San Francisco and the University of Washington, Seattle, Washington. 4
- Brainard, William C.** 1967. “Uncertainty and the Effectiveness of Policy.” *American Economic Review*, 57(2): 411–425. 5
- Coibion, Olivier, and Yuriy Gorodnichenko.** 2012. “Why Are Target Interest Rate Changes So Persistent?” *American Economic Journal: Macroeconomics*, 4(4): 126–62. 5
- Cukierman, Alex.** 1991. “Why Does the Fed Smooth Interest Rates?” In *Monetary Policy on the 75th Anniversary of the Federal Reserve System*. 111–147. Springer. 5
- Geist, Matthieu, and Olivier Pietquin.** 2011. “Kalman Filtering & Colored Noises: the (Autoregressive) Moving-Average Case.” 1–4. 36
- Klein, Michael W., and Jay C. Shambaugh.** 2008. “The dynamics of exchange rate regimes: Fixes, floats, and flips.” *Journal of International Economics*, 75(1): 70–92. 14, 15
- Levin, Andrew, Volker Wieland, and John C. Williams.** 2003. “The Performance of Forecast-Based Monetary Policy Rules Under Model Uncertainty.” *American Economic Review*, 93(3): 622–645. 5
- Milani, Fabio.** 2007. “Expectations, learning and macroeconomic persistence.” *Journal of Monetary Economics*, 54(7): 2065–2082. 5
- Onatski, Alexei, and Noah Williams.** 2003. “Modeling Model Uncertainty.” *Journal of the European Economic Association*, 1(5): 1087–1122. 11
- Orphanides, Athanasios.** 2001. “Monetary Policy Rules Based on Real-Time Data.” *American Economic Review*, 91(4): 964–985. 11
- Orphanides, Athanasios.** 2003a. “Historical Monetary Policy Analysis and the Taylor Rule.” *Journal of Monetary Economics*, 50(5): 983–1022. 5

- Orphanides, Athanasios.** 2003*b*. “Monetary Policy Evaluation with Noisy Information.” *Journal of Monetary Economics*, 50(3): 605–631. 4
- Rotemberg, Julio J., and Michael Woodford.** 1999. “Interest Rate Rules in an Estimated Sticky Price Model.” *Monetary Policy Rules*, 57–126. University of Chicago Press. 5
- Rudebusch, Glenn, and Lars E.O. Svensson.** 1999. “Policy Rules for Inflation Targeting.” *Monetary Policy Rules*, 203–262. University of Chicago Press. 5, 17, 21, 33
- Rudebusch, Glenn D.** 2001. “Is the Fed Too Timid? Monetary Policy in an Uncertain World.” *Review of Economics and Statistics*, 83(2): 203–217. 4, 5, 10, 11
- Rudebusch, Glenn D.** 2006. “Monetary Policy Inertia: Fact or Fiction?” *International Journal of Central Banking*, 2(4): 85–135. 4
- Sack, Brian.** 1998. “Uncertainty, Learning, and Gradual Monetary Policy.” Board of Governors of the Federal Reserve System Discussion Paper 1998-34. 5
- Sack, Brian.** 2000. “Does the Fed Act Gradually? A VAR Analysis.” *Journal of Monetary Economics*, 46(1): 229–256. 5
- Shambaugh, Jay C.** 2004. “The Effect of Fixed Exchange Rates on Monetary Policy*.” *The Quarterly Journal of Economics*, 119(1): 301–352. 14, 15
- Söderström, Ulf.** 2002. “Monetary Policy with Uncertain Parameters.” *Scandinavian Journal of Economics*, 104(1): 125–145. 5
- Stein, Jeremy C, and Adi Sunderam.** 2015. “Gradualism in Monetary Policy: A Time-Consistency Problem?” National Bureau of Economic Research Working Paper 21569. 5
- Woodford, Michael.** 1999. “Optimal Monetary Policy Inertia.” *The Manchester School*, 67: 1–35. 5
- Woodford, Michael.** 2003. “Optimal Interest-Rate Smoothing.” *The Review of Economic Studies*, 70(4): 861–886. 5

Appendices

A Kalman filter with MA(1) noises

Standard Kalman filter assumes zero mean Gaussian white noise for w_t . In this paper, I introduce an auxiliary random process so that w_t follows MA(1) within the Kalman filter.²¹ The MA(1) noise defined in equations (23) and (24) is equivalent to the following:

$$\begin{bmatrix} w_t \\ \phi_t \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{2,2} & \boldsymbol{\theta} \\ \mathbf{0}_{2,2} & \mathbf{0}_{2,2} \end{bmatrix} \begin{bmatrix} w_{t-1} \\ \phi_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_t \end{bmatrix} \quad (42)$$

where

$$\phi_t = \begin{bmatrix} \phi_t^\pi \\ \phi_t^y \end{bmatrix}, \boldsymbol{\theta} = \begin{bmatrix} \theta^\pi & 0 \\ 0 & \theta^y \end{bmatrix}, \boldsymbol{\xi}_t = \begin{bmatrix} \epsilon_t^\pi \\ \epsilon_t^y \end{bmatrix}.$$

ϕ_t is the auxiliary vector to transfer past white noise to current period, $\boldsymbol{\theta}$ is the matrix of MA(1) parameters, and $\boldsymbol{\xi}_t$ is the vector of white noises. The upper block ($w_t = \boldsymbol{\theta}\phi_{t-1} + \boldsymbol{\xi}_t$) and the lower block ($\phi_t = \boldsymbol{\xi}_t$) together produce $w_t = \boldsymbol{\xi}_t + \boldsymbol{\theta}\boldsymbol{\xi}_{t-1}$.

Given the noise formulation (42), an equivalent to state-space (27) and (28) with MA(1) observation noise can be proposed:

$$\begin{bmatrix} X_{t+1} \\ w_{t+1} \\ \phi_{t+1} \end{bmatrix} = \begin{bmatrix} A & \mathbf{0}_{10,2} & \mathbf{0}_{10,2} \\ \mathbf{0}_{2,10} & \mathbf{0}_{2,2} & \boldsymbol{\theta} \\ \mathbf{0}_{2,10} & \mathbf{0}_{2,2} & \mathbf{0}_{2,2} \end{bmatrix} \begin{bmatrix} X_t \\ w_t \\ \phi_t \end{bmatrix} + \begin{bmatrix} B \\ \mathbf{0}_{2,1} \\ \mathbf{0}_{2,1} \end{bmatrix} i_t + \begin{bmatrix} \boldsymbol{\nu}_{t+1} \\ \boldsymbol{\xi}_{t+1} \\ \boldsymbol{\xi}_{t+1} \end{bmatrix} \quad (43)$$

$$Z_t = \begin{bmatrix} C & \mathbf{I}_{2,2} & \mathbf{0}_{2,2} \end{bmatrix} \begin{bmatrix} X_t \\ w_t \\ \phi_t \end{bmatrix} \quad (44)$$

Therefore, we can rewrite the equation (27) and (28) as

$$X'_{t+1} = A'X'_t + B'i_t + \boldsymbol{\nu}'_{t+1} \quad (45)$$

$$Z'_t = C'X'_t \quad (46)$$

²¹Please refer to [Geist and Pietquin \(2011\)](#) for more detailed explanations on autoregressive and moving-average noise in Kalman filter.

where

$$X'_t = \begin{bmatrix} X_t \\ \mathbf{w}_t \\ \phi_t \end{bmatrix}, A' = \begin{bmatrix} A & \mathbf{0}_{10,2} & \mathbf{0}_{10,2} \\ \mathbf{0}_{2,10} & \mathbf{0}_{2,2} & \boldsymbol{\theta} \\ \mathbf{0}_{2,10} & \mathbf{0}_{2,2} & \mathbf{0}_{2,2} \end{bmatrix}, B' = \begin{bmatrix} B \\ \mathbf{0}_{2,1} \\ \mathbf{0}_{2,1} \end{bmatrix}, \boldsymbol{\nu}'_t = \begin{bmatrix} \boldsymbol{\nu}_t \\ \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_t \end{bmatrix}, C' = \begin{bmatrix} C \\ \mathbf{I}_{2,2} \\ \mathbf{0}_{2,2} \end{bmatrix}^T$$

Optimal Kalman gain K' (14×2 matrix) and predicted estimate covariance $P'_{t|t-1}$ (14×14 matrix) are specified as

$$\begin{aligned} K' &= P'_{t|t-1} C'^T (C' P'_{t|t-1} C'^T)^{-1} \\ P'_{t|t-1} &= A' (P'_{t|t-1} - K' C' P'_{t|t-1}) A'^T + V'_\nu \end{aligned} \quad (47)$$

where V'_ν (14×14 matrix) is variance-covariance matrices of $\boldsymbol{\nu}'_t$.

The central bank's optimal inference $X'_{t|t}$ and forecast $X'_{t+1|t}$ are in recursive form as

$$\begin{aligned} X'_{t|t} &= A' X'_{t-1|t-1} + B' i_{t-1} + K' (Z'_t - Z'_{t|t-1}) \\ &= (I - K' C') A' X'_{t-1|t-1} + (I - K' C') B' i_{t-1} + K' Z'_t \end{aligned} \quad (48)$$

$$X'_{t+1|t} = A' X'_{t|t} + B' i_t \quad (49)$$

and the estimated values of $X'_{t|t}$ and $X'_{t+1|t}$ are entered as $X_{t|t}$ and $X_{t+1|t}$ in equation (30) and (31) after eliminating the auxiliary variables.

B Additional Tables and Figures

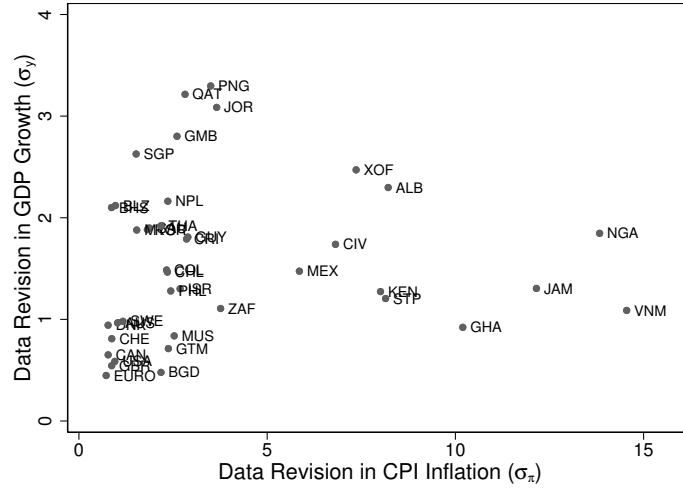
Table 1: Example of historical WEO data (United States)

	WEO Releases (Real GDP Growth)					
	S1990	F1990	S1991	F1991	S1992	F1992
1988	4.06	4.08
1989	2.99	2.96	3.26	3.31	.	.
1990	2.27	1.98	2.08	2.20	2.19	2.26
1991	3.09	2.44	1.19	0.86	-0.33	0.15
1992	3.30	3.42	2.93	2.79	1.42	1.10
1993	3.33	3.45	3.42	3.36	3.57	3.14
1994	3.32	3.42	3.46	3.41	3.75	3.55
1995	3.34	3.40	3.39	3.29	3.77	3.73
1996	.	.	3.39	3.29	3.62	3.74
1997	3.57	3.67

	WEO Releases (CPI Inflation)					
	S1990	F1990	S1991	F1991	S1992	F1992
1988	4.09	4.07
1989	4.80	4.77	4.82	4.82	.	.
1990	4.56	5.08	5.28	5.40	5.41	5.40
1991	4.19	4.51	4.90	4.53	4.23	4.28
1992	4.20	4.20	4.00	4.05	3.11	3.11
1993	4.20	4.20	3.90	3.80	3.07	3.13
1994	4.20	4.20	3.70	3.60	3.00	3.00
1995	4.20	4.20	3.60	3.40	2.90	2.90
1996	.	.	3.50	3.30	2.75	2.80
1997	2.60	2.80

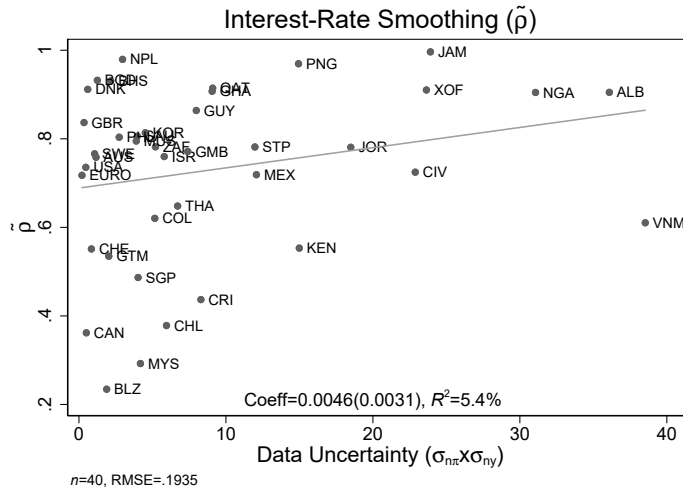
Note: This example includes the WEO issues from Spring 1990 to Fall 1992. Values are in percentage change from previous years. Sample data span from Spring 1990 to Fall 2008 for each country.

Figure 12: Data uncertainty in CPI and GDP countries (raw standard deviation)



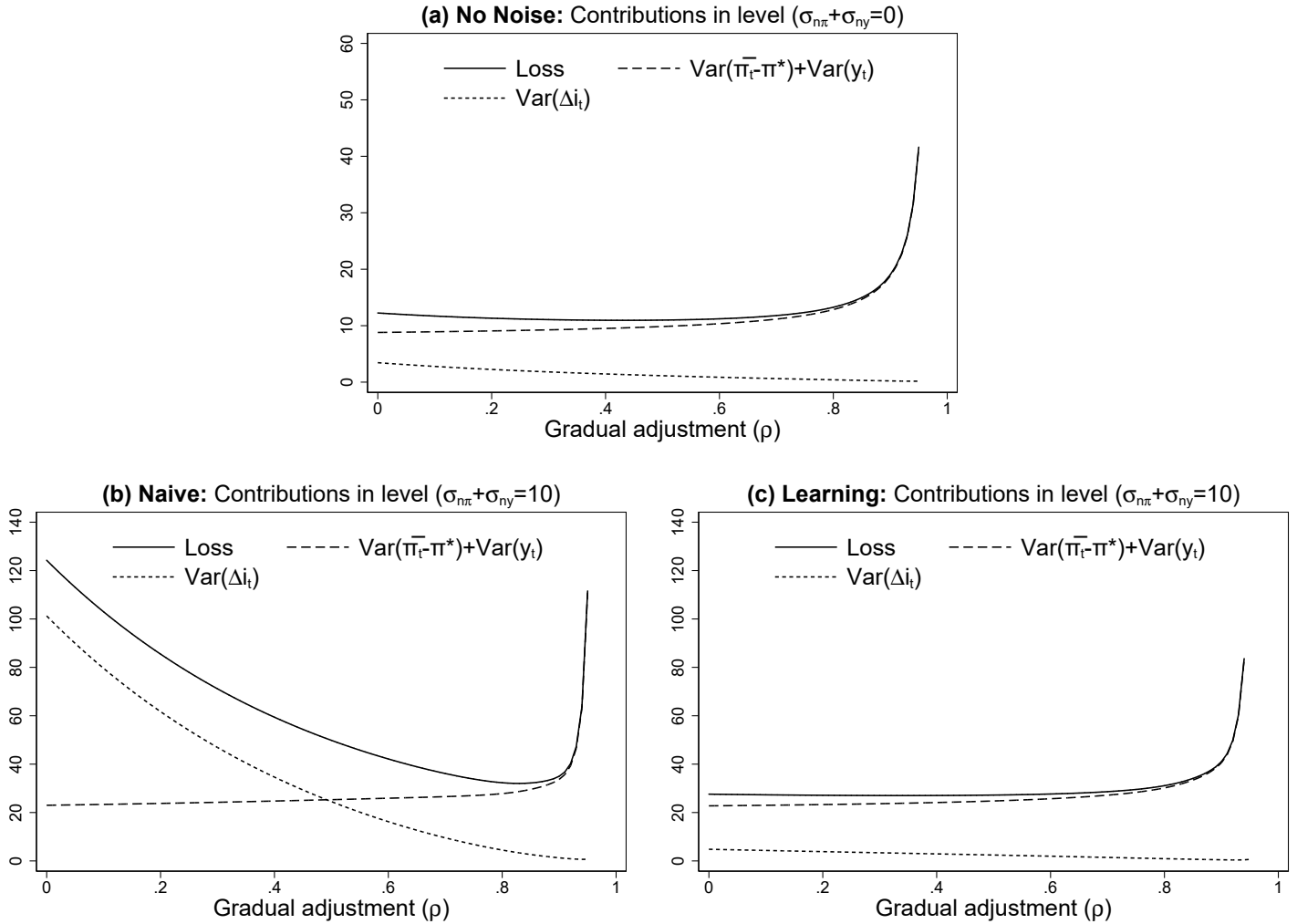
Note: Raw standard deviations are reported for data revisions in CPI inflation and GDP growth. On the contrary, figure 2 (a) reports standard errors implied by MA(1) noise process. Both provide the comparatively similar amount of data revisions with minor differences in magnitude.

Figure 13: Cross-country comparison of interest-rate smoothing (product measure)



Note: An alternative measure of data uncertainty is considered: a product of standard errors ($\sigma_{n\pi} \times \sigma_{ny}$) instead of the sum of standard errors ($\sigma_{n\pi} + \sigma_{ny}$). The scale of data uncertainty increases from 20 to 40, and the overall pattern is similar to that in figure 2 (b).

Figure 14: Loss decomposition (absolute contributions)



Note: Figure 14 reports the loss decomposition under different monetary policy and level of data uncertainty. It shows the contribution of $\text{Var}[\bar{\pi}_t - \pi^*] + \text{Var}[y_t]$ and $\text{Var}[\Delta i_t]$ in levels and their associated loss with regard to the degree of gradual adjustment (ρ). Note that $\text{Var}[\bar{\pi}_t - \pi^*]$ and $\text{Var}[y_t]$ increase in ρ and $\text{Var}[\Delta i_t]$ decreases in ρ . $\text{Var}[\bar{\pi}_t - \pi^*]$ and $\text{Var}[y_t]$ diverge to infinity if ρ is close to 1. Panel (a) reports those when there is no data uncertainty, and they are same under both monetary polices. Panel (b) reports those in face of data uncertainty ($(\sigma_{n\pi} + \sigma_{ny}) = 10$) under naive policy, and panel (c) reports those in face of data uncertainty ($(\sigma_{n\pi} + \sigma_{ny}) = 10$) under learning policy. $\text{Var}[\Delta i_t]$ curve becomes much steeper in the face of data uncertainty under naive policy. $\text{Var}[\Delta i_t]$ curve becomes relatively flatter in the face of data uncertainty under learning policy. Figure 15 reports the relative contributions of each loss components for a better observation of the change in $\text{Var}[\Delta i_t]$ from panel (a) to (c). Figure 7 provides diagrams that summarize these changes.

Figure 15: Loss decomposition (relative contributions)

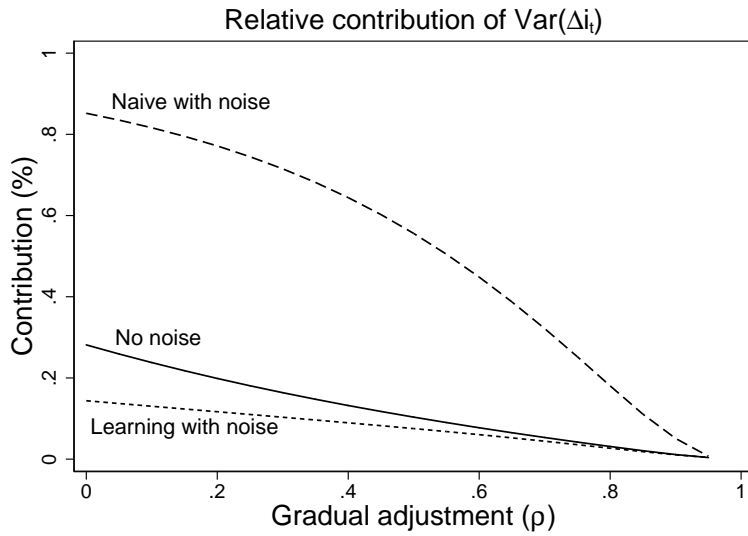


Figure 15 reports the relative contributions of each loss components for a better observation of the change in $Var[\Delta i_t]$ under naive and learning policies. The relative contribution curve of $Var[\Delta i_t]$ becomes much steeper in the face of data uncertainty under the naive policy and flatter in the face of data uncertainty under the learning policy.