IMF Working Paper

Testing the Transparency Benefits of Inflation Targeting: Evidence from Private Sector Forecasts

Christopher Crowe

IMF Working Paper

Research Department

**Testing the Transparency Benefits of Inflation Targeting:
Evidence from Private Sector Forecasts**

Prepared by Christopher Crowe[1]

Authorized for distribution by Paolo Mauro

December 2006

Abstract

I test whether inflation targeting (IT) enhances transparency using inflation forecast data for 11 IT adoption countries. IT adoption promotes convergence in forecast errors, suggesting that it enhances transparency. This effect is robust to dropping observations, is strengthened by using instrumental variable estimation to eliminate mean-reversion, and is absent in placebo regressions (where IT adoption is shifted by a year). This result supports Morris and Shin's (2002) contention that better public information is most beneficial for forecasters with bad private information. However, it does not support their hypothesis that better public information could make private forecasts less accurate.

Author's E-Mail Address: ccrowe@imf.org

Contents

I. INTRODUCTION

> "The most important distinguishing characteristic of inflation target regimes
> is the emphasis they place on transparency and accountability."
> —Mervyn King, 1997

The above comment by Mervyn King, then Deputy Governor and now Governor of the Bank of England, represents the majority view among policymakers and economists over the transparency-enhancing effects of inflation targeting (IT). However, few studies up till now have investigated whether IT confers these benefits, and those that have done have tended to focus on subjective or qualitative measures of transparency or on a relatively narrow set of countries. This study attempts to fill this gap.

As befits a discussion of transparency, let us first be transparent about what we take the term to mean. Following Geraats (2002), transparency is defined for the purposes of this paper as "the removal of information asymmetries." Geraats catalogues five areas of monetary policymaking where transparency could effect outcomes: the central bank's objectives and its institutional relationship with the rest of government, the publication of data and forecasts, internal decisionmaking, communication and explanation of policy changes and details of the implementation of policy; these in turn give rise to five elements of transparency (*political, economic, procedural, policy* and *operational*). The introduction of IT is widely held to increase political transparency, by forcing the central bank to specify both the variable that it is targeting (some measure of consumer price inflation) and a precise numerical target for the targeted variable, as well as—in many cases—delineating more clearly the division of responsibilities between the bank and the political authorities (King, 1997; Eijffinger and Geraats, 2006). The communication strategy accompanying the targeting regime typically includes enhanced policy analysis, more openness about internal policy deliberations and greater explanation of policy decisions, resulting in enhanced procedural, policy, and operational transparency (Berg, 2005). Finally, this communication strategy usually involves a more explicit and public discussion of the forecasts underlying policy decisions, including attendant risks and underlying assumptions, which tends to increase economic transparency (Roger and Stone, 2005).

The ultimate goals of enhanced transparency, albeit ones that are difficult to quantify, are to increase the effectiveness of monetary policy and to make otherwise independent institutions more democratically accountable (Blinder and others, 2001). A narrower benefit of enhanced transparency is that, with a fuller understanding of the central bank's objectives and decision-making procedures and access to the central bank's forecasts and analysis, the private sector's forecasts of variables over which the central bank has some influence should become more accurate. The current paper's objective is to assess whether this specific and clearly measurable benefit of IT exists: the paper tests the joint hypothesis that (a) transparency improves the accuracy of private sector forecasts and (b) IT enhances transparency.

Both hypotheses seem plausible ex ante, and the balance of the theoretical literature is in their favor. Geraats (2002) and Carpenter (2004) provide surveys of the existing literature

on hypothesis (a), regarding the potential benefits of greater transparency. Early theoretical contributions (of which Cukierman and Meltzer, 1986, is the most widely cited) generally adopt the Barro-Gordon (1983) or Kydland-Prescott (1977) model in which the authorities are able to temporarily boost output by creating unanticipated monetary expansions, at the potential cost of imparting an inflationary bias to policymaking. In this environment, greater transparency can reduce the inflation bias, but it can also limit the central bank's ability to offset output surprises. This cost of transparency generally requires the existence of some distortion (in addition to the information assymetry between the policymaker and the private sector). Later contributions have tended to downplay the ability of the central bank to offset these distortions, stressing instead the primacy of long-term policy goals (particularly price stability), which can be better achieved if the policymaker enjoys credibility. King (1997) echoes other authors in arguing that enhanced transparency helps to achieve this.

For the current paper, a key contribution to this debate is that of Morris and Shin (2002; hereafter, MS), who dissent from the apparent consensus concerning the benefits of greater transparency. Specifically, if the private sector attempts to second-guess itself in the manner of Keynes's (1936) "beauty contest," then public information, acting as a focal point for "beliefs about beliefs," can crowd out high-quality private information and make private sector forecasts *more variable*, not less.[2] As outlined in Appendix I, they show that (i) enhanced transparency is most likely to benefit forecasters whose own forecasts are less accurate, while (ii) the best forecasters can, in principle, be harmed by more transparent public information.

To capture these predictions in a simple framework, the effect of enhanced transparency on forecast errors V can be thought of as the sum of two effects: a constant effect for all forecasters (irrespective of forecasting ability), and an interaction effect that falls with forecast accuracy. Proxying for (the inverse of) forecasters' ability using the initial forecast error V_0, then the effect of better public information on forecast accuracy can be modeled as:

$$\Delta V = b_0 - b_1 V_0. \tag{1}$$

MS's prediction (i) corresponds to the hypothesis that $-b_1 < 0$ (the worst forecasters —those with higher initial forecast errors—experience the greatest improvement in forecast accuracy) while prediction (ii) corresponds to the hypothesis that $b_0 > 0$ (the best forecasters—those with $V_0 \longrightarrow 0$—see their forecast errors increase when public information improves). Section II outlines a simple signal extraction model (a special case of MS with the weight on the beauty contest motivation set at zero) in which equation (1) is explicitly derived. While $-b_1 < 0$ is consistent with the simple model, $b_0 > 0$ is particular to MS and relies on some additional parameter restrictions as outlined in Appendix I. Both hypotheses are tested in section III in the context of IT adoption, and therefore shed some light on whether MS's concerns over transparency are likely to hold. To preview the results, there is strong evidence for the first hypothesis but no evidence for the second, suggesting that

---

[2]Svensson (2006) shows that MS's result, that publishing central bank forecasts can make private sector forecasts more variable, relies on unlikely parameter values; in the more general case the central bank can publish and not be damned. This is discussed further in Appendix I.

while IT enhances transparency and is particularly helpful for those with poor private information, it is unlikely to harm those whose private signals are already precise.

Chortareas, Stasavage, and Sterne (2002) present some empirical evidence relating to these questions. They find that greater transparency, in the form of more public prominence for the central bank's forecasts, is associated with lower inflation in a diverse cross-section of countries. This result perhaps speaks more to the credibility benefits of transparency rather than directly to the benefits in terms of more accurate private sector forecasts.[3] As MS argue, greater transparency over forecasts is most beneficial when the central bank's forecast is more accurate than those of the private sector. Romer and Romer (2000) analyze whether unpublished central bank forecasts (specifically, by the Federal Reserve) are actually superior to the professional forecasts of the private sector, and they find persuasive evidence that this is the case. In fact, the Fed's unpublished forecasts are so good that if the private sector forecasters had access to them they would place no weight on their own forecasts.[4] This points to significant potential benefits to greater transparency by the Fed, although the authors note that if the forecasts were published (and expected to be published), they might become less accurate.

Cross-country comparison studies have tended to criticize the Fed for lagging behind in the general movement toward greater transparency. For instance, Eijffinger and Geraats (2006) note that the Fed's political transparency is weaker than that of comparable institutions (particularly inflation targeters). However, they note that its policy transparency compares favorably with IT institutions. Swanson (2004) provides some evidence in support of this: financial market forecasts of policy interest rates in the United States have improved markedly over time, in step with transparency-enhancing changes to the Fed's communications strategy and policymaking.[5]

Opinion on hypothesis (b) is less divided. Most commentators agree that a key characteristic of IT regimes (for many, *the* key difference) is their greater degree of transparency (Bernanke and others (1999); Faust and Henderson (2004); King (1997); Mishkin and Schmidt-Hebbel (2001); Svensson (1999)). Svensson's contention, mirroring Mervyn King's sentiments, that "inflation targeting regimes are characterized by a high

---

[3]The transparency measure is derived from self-reported information on central bank governance in a wide-ranging survey of central banks (Fry and others, 2000); to this extent it may be less contaminated by tautological reasoning than the transparency indices employed elsewhere (although additional biases may arise from using survey data).

[4]This is not to say that private sector forecasts are themselves bad: Ang, Bakaert, and Wai (2005) find that, in the United States, inflation forecasts from surveys (from both professional and nonprofessional forecasters) are better predictors of future inflation than model-based forecasts or implied forward inflation from financial market data.

[5]A related issue, particularly relevant to discussions of transparency at the Fed, is how the publication of voting records and policy discussion transcripts (with a delay) affects transparency and the incentives facing individual committee members (see Meade, 2006).

degree of transparency and accountability" is representative of the received wisdom.[6] This academic consensus also seems to be reflected in press accounts of market sentiment. One potential transparency benefit of IT is that it depersonalizes monetary policy. Because IT is more transparent about the intentions or goals of policy, the private sector understands that the central bank's objectives do not depend on the preferences of the central bank's management, and as a result is less spooked by the appointment of new senior officials (King, 1997). And because actions and forecasts are also more transparent, the private sector can more easily verify that policy does indeed reflect institutional, not personal, preferences (Svensson, 1999). It is noticable from press discussions of personnel changes that the focus on individual preferences and character does seem more pronounced for those central banks, such as the Fed, that have not adopted IT.[7]

Ultimately, whether IT enhances transparency is an empirical question, and here evidence is rather limited. Eijffinger and Geraats (2006) attempt to quantify central bank transparency by analyzing the monetary authorities' relationship with the political authorities and how policy is formulated and communicated. Their analysis of nine major central banks concludes that "the most transparent central banks ... are all inflation targeters." Roger and Stone (2005), who measure transparency according to central banks' adherence to the IMF's Code of Monetary and Financial Policy Transparency, come to the same conclusion. The validity of these results relies on the suitability of the subjective transparency measure employed: as with any subjective index, there is a risk of tautological reasoning biasing the results toward finding an effect.

Several papers have investigated the behavior of inflation under IT (Ball and Sheridan, 2004; Kuttner and Posen, 1999; Petursson, 2004; Vega and Winkelried, 2005) with mixed results; a related question, and one that is more interesting in the current context, is IT's effect on inflation expectations. Gurkaynak, Levin, and Swanson (2005) use bond yields in the United Kingdom, Sweden and the United States to back out expected inflation and infer the effects of the introduction of IT in the two former countries. They find that IT

---

[6]However, Friedman (2004) argues that IT in fact lacks transparency because it tends to be unclear what weight the central bank places on minimizing output variations (although it is not obvious that clarity is any greater under alternative policy regimes). Friedman's criticism, if valid, would lead one to question the view that IT enhances political transparency; the view that it enhances the four other aspects of transparency would seem still to hold. Moreover, a preliminary analysis of IT's effect on output growth forecasts suggests that IT adoption has little or no effect on forecast accuracy (results available from the author); this would tend to suggest that Friedman's concerns are unwarranted, since if the central bank put less effort into minimizing output fluctuations under IT, they would presumably become harder to forecast.

[7]As an example, consider market reaction to the appointment of a new Federal Reserve Chairman in 2006. It seems clear from news accounts of the transition that the change in personnel increased uncertainty for financial market participants in the absence of formal policy targets. Hence, the *Wall Street Journal* (January 30, 2006) predicted that "bonds, and to a lesser extent stocks, likely will hit a patch of turbulence in coming weeks as markets try to figure out how much and which way the Federal Reserve's incoming chairman, Ben Bernanke, will prod the economy." Contrast this with a press account of the appointment of Alan Bollard as Governor of the Reserve Bank of New Zealand (RBNZ; the first central bank to formally adopt IT). The *Dominion Post* (August 23, 2002) analyzed the implications for policy of Dr. Bollard's appointment and of the concurrent decision over whether to extend the inflation targeting agreement between the RBNZ and the authorities, concluding that "the real key to the central bank's future attitude toward tackling inflation will be the new policy targets agreement."

makes long-term inflation expectations less responsive to economic "news", suggesting that IT helps to anchor long-term inflation expectations. This effect cannot be ascribed necessarily to greater transparency under IT: enhanced credibility—distinct from but related to transparency—may be the explanation. Corbo, Landerretche and Schmidt-Hebbel (2001) analyze model-derived inflation forecasts for a range of advanced and emerging market economies, and find a drop in forecast errors among countries that adopt IT, although the fall seems to predate the adoption of IT in many cases and may reflect other changes to the policy environment. Johnson (2002) comes closest to the current paper in terms of research question and methodology. He tests the credibility and transparency effects of IT, using similar data on private sector inflation forecasts, in a panel of eleven industrial countries including five inflation targeters. He finds some evidence for credibility benefits of IT (expected inflation falls more in the targeters than the non-targeters) but no evidence for transparency benefits (neither the forecasts' variability nor their absolute error are reduced by IT). This suggests that MS's concerns may be relevant, although this result may also be due do methodological difficulties, as discussed in section III.

Section II outlines a simple signal extraction model with public and private information to motivate the discussion and provide some predictions This model is a special case of MS—Appendix I provides a brief summary of the main predictions in the context of the more complex MS model. Section III desribes the empirical strategy, the data used and the main results—there is a more detailed description of the matching methodology employed in Appendix II. Section IV offers conclusions.

II.   Theoretical Framework

I motivate the empirical analysis via a simple signal extraction model.[8] Agents ("forecasters") seek to minimize the squared error of their inflation forecast f_i around the actual inflation rate π (so that the marginal disutility of the forecast error increases with the magnitude of the error):

$$u_i\left(f_i, \pi\right) \equiv -\left(f_i - \pi\right)^2. \tag{2}$$

The private sector agents observe the central bank's public signal (a combination of its public forecasts, statements and analysis), π_C, and also observe their own private signal π_i. Each signal is noisy:

$$
\begin{aligned}
\pi_C &= \pi + \eta \\
\pi_i &= \pi + \varepsilon
\end{aligned}
\tag{3}
$$

---

[8]Apendix I derives the relevant comparative statics results under MS's more general model of strategic behavior in response to public and private information. The analysis replicates that of MS, focusing more explicitly on the predictions most relevant for the current paper. The simple signal extraction model presented here is in fact a special case of MS with the strategic element shut off by setting their r parameter to zero.

and the precision of the public and private signals is denoted, respectively, as:

$$\alpha = \frac{1}{\sigma_\eta^2} \tag{4}$$

$$\beta = \frac{1}{\sigma_\varepsilon^2}.$$

Agents therefore optimally weight the two signals according to their relative precision:

$$f_i^* = \frac{\alpha \pi_C + \beta \pi_i}{\alpha + \beta}. \tag{5}$$

Then the expected mean square forecast error (a measure of the forecast inaccuracy) is given by:

$$\widetilde{V} \equiv E\left[(f_i - \pi)^2\right] = \frac{1}{(\alpha + \beta)}. \tag{6}$$

I introduce some identifying assumptions: (a) that the precision of the private signals is constant, for each forecaster i, over time; and (b) and that the precision of the public signal depends only on some country-specific factor and whether the central bank of country j has adopted inflation targeting ($IT = \{0, 1\}$):

$$\alpha_t^j = \alpha^j(IT) \tag{7}$$

$$\beta_t^i = \beta^i$$

Hence the statement that IT improves transparency is equivalent to the condition that $\alpha^j(1) > \alpha^j(0)$. The forecast error for a typical forecaster in a non-IT ($IT = 0$) central bank is then given by:

$$\left(\widetilde{V}^{ij} \mid IT = 0\right) \equiv \widetilde{V}_0^{ij} = \frac{1}{\left(\alpha^j(0) + \beta^i\right)} \tag{8}$$

and hence:

$$\frac{\partial}{\partial \alpha^j} \widetilde{V}^{ij} = -\frac{1}{\left(\alpha^j(0) + \beta\right)^2} = -\left(\widetilde{V}_0^{ij}\right)^2 < 0 \tag{9}$$

$$\frac{\partial^2 \widetilde{V}ij}{\partial \alpha^j \partial \widetilde{V}_0^{ij}} = -2\widetilde{V}_0^{ij} < 0 \tag{10}$$

Linearizing the interaction effect around V_0^{ij} then gives the following approximation for the effect of IT on forecast errors:

$$\Delta \widetilde{V}^{ij} \equiv \widetilde{V}_1^{ij} - \widetilde{V}_0^{ij} \simeq b_0 - b_1 \widetilde{V}_0^{ij} \tag{11}$$

$$-b_1 < 0$$

This formulation suggests the use of a "difference-in-difference" approach: that is, comparing the behavior of forecasts, across forecasters, in countries that adopted IT (the "treatment" group) with the behavior of forecasters in similar countries that did not adopt IT (the "control" group). The next section presents the empirical strategy.

III.   Empirical Strategy and Results

A.   Data and Methodology

Before taking equation (11) to the data, I provide a description of the data itself, since the nature and availability of data largely dictate the choice of empirical strategy. I use the *Consensus Forecasts* dataset, which comprises a panel of private sector "current year" and "next year" forecasts of several key macroeconomic variables. For this study I focus on the forecasts of inflation, since it is for inflation expectations that IT's transparency benefits are usually held to be strongest. The country coverage expands over time, from a small number of industrial countries at the end of 1989 to a large cross-section of industrial and emerging market economies by 2005. Some countries that adopted IT were not in the sample at the time of adoption (even if they later joined the sample).[9] For our purposes, eleven IT-adoption episodes are covered by the dataset: four industrial countries (Australia, Canada, Norway, and the United Kingdom) and seven emerging markets (Brazil, Chile, Colombia, Korea, Mexico, Peru, and Thailand). I date the adoption of IT to a specific month and year from Roger and Stone (2005), which seems, for most countries, to represent a broad consensus view.[10]

For our purposes the "next year" forecasts are most useful (Johnson, 2002, who uses the same data for some of the analysis, makes the same decision: the "current year" forecasts tend to vary little across forecasters, particularly toward the end of the year, for obvious reasons).[11] For each country there is a panel of forecasters whose composition changes somewhat over time as individual forecasters enter or drop out of the survey. Forecasts are monthly or, for some countries, bimonthly. In order to control for composition effects and exploit within-country variation in forecaster quality (to test the interaction effect captured by the parameter b_1), I focus on individual forecasters.[12] I identify a 24-month window spread equally either side of the adoption of IT: it makes sense to focus on a relatively narrow window to exploit the monthly nature of the data and identify more sharply the effect of IT.[13] Then I use the average (per-forecaster) absolute forecast error (the absolute difference between the "next year" inflation forecast and actual (annual) inflation next year, taken from the IMF's *International Financial Statistics*) in the "before" and "after" portions of the window as proxies for $\widetilde{V}_0^{ij}$ and $\widetilde{V}_1^{ij}$, respectively. To borrow the terminology from incidence analysis, the 166 forecasters for whom we have "before" and "after" data in

---

[9] These include the Czech Republic, Hungary, New Zealand, Poland, and Sweden.

[10] The dating is extremely clear for some countries; for others there is some controversy as to the precise month that IT was adopted.

[11] The forecasts, although they are collected monthly or bimonthly, refer to calendar years rather than a 12-month-ahead moving window. As an example, the "next year" forecasts from January 1991 through December 1991 are all for same 12-month period ending in December 1992.

[12] Forecasters are identified in the survey by their name: there are some minor name changes (some genuine, some apparently due to spelling errors), which complicate attempts to correctly match observations to each individual forecaster. I attempt to overcome this via an algorithm that identifies individual forecasters.

[13] Transparency benefits are likely to occur relatively quickly, compared with credibility benefits or effects on actual variables. Annual data may be too coarse to pick up any effects: this could help to explain Johnson's (2002) negative findings.

the window around IT's adoption, in the 11 countries that adopt IT, form our "treatment" group.

It is important to recognize that V^{ij}, our empirical counterpart for $\widetilde{V}^{ij}$, will be contaminated by idiosyncratic time-varying shocks to forecasters' accuracy as well as by classical measurement error. If we capture these shocks by the linear error term e_t^{ij}, then the empirical counterpart of equation (11) is given by:

$$\Delta V^{ij} \equiv V_1^{ij} - V_0^{ij} = \begin{cases} b_0 - b_{10}^{ij} V^{ij} + \Delta e_t^{ij} \mid IT = 1 \\ \Delta e_t^{ij} \mid IT = 0 \end{cases}. \tag{12}$$

To test equation (12) I estimate the following regression:

$$\Delta V^{ij} = b_0 + b_{0T} D_T^{ij} - V_0^{ij} \left(b_1 + b_{1T} D_T^{ij} \right) + u^{ij}, \tag{13}$$

where D_T is a dummy variables for the "treatment" (IT adoption).

From equation (12), the error term u^{ij} will include the change in the ideosyncratic shocks to forecast accuracy, Δe^{ij}. Since $\Delta e^{ij} \equiv e_1^{ij} - e_0^{ij}$ and $V_0^{ij} = \frac{1}{(\alpha^j(0)+\beta^i)} + e_0^{ij}$, then $cov\left(V_0^{ij}, \Delta V^{ij}\right) < 0$: the data will exhibit mean reversion. Among the worst forecasters in period 0 will be those whose forecasts in that period were particularly poor quality, compared with their average performance, and these forecasters will naturally experience an improvement in their performance in period 1. Furthermore, $cov\left(V_0^{ij}, u^{ij}\right) < 0$, so that the estimate of the coefficient on V_0^{ij} will be negatively biased. I include the term $-b_1 V_0^{ij}$ to control for mean reversion, which should remove any bias from our estimate of the effect of IT itself (captured by the term $-b_{1T} D_T V_0^{ij}$). This strategy mimics that of Ball and Sheridan (2004). I test the model via ordinary least squares (OLS) estimation of equation (13).

However, this solution is at best a partial one. The econometric problem posed by mean reversion is endogeneity: i.e., correlation between one of the regressors (V_0^{ij}) and the error term. A superior solution is therefore offered by adopting an instrumental variables (IV) estimation procedure. Potential instruments should be correlated with the "fundamental" component of forecaster ability $\left(\frac{1}{(\alpha^j(0)+\beta^i)}\right)$ but uncorrelated with the transient shocks to forecast accuracy e_0^{ij} that are generating the endogeneity problem.

I identify two such instruments for V_0^{ij}. The first instrument $V_{0,g}^{ij}$ is the direct counterpart of V_0^{ij}, but relating to forecasters' estimate of GDP growth rather than inflation.[14] Assuming that ideosyncratic shocks to forecast accuracy for inflation and GDP growth are orthogonal but that forecasters' fundamental ability is reflected in the accuracy of both forecasts, then

---

[14] Actual (next year) GDP growth data are taken from the IMF's *International Financial Statistics*. Some gaps in the data are filled in using data from the IMF's *World Economic Outlook* database.

this should fulfill the relevancy and exogeneity requirements of a "good" instrument.[15]

The second instrument is based on the observation that higher inflation also tends to be more variable, and hence likely to be harder to forecast. Specifically, I use the expected inflation rate for the one-year period prior to the IT adoption date, f^{ij}. Using expected rather than actual inflation eliminates the impact of unexpected inflationary shocks that are likely to be correlated with the transient component of forecast errors.[16]

These two variables $\left\{V_{0,g}^{ij}, f^{ij}\right\}$ constitute suitable instruments for V_0^{ij}. Then, since the "treatment" dummy D_T is assumed exogenous (conditional on the matching of a control group of observations with the treatment observations as described below), suitable instruments for $D_T V_0^{ij}$ are $\left\{D_T V_{0,g}^{ij}, D_T f_i\right\}$.

B.   Matching Treatment and Control Groups

As emphasized by Besley and Case (2000), "the quality of difference-in-differences estimation is crucially dependent on the quality of the control group chosen." I match forecasters in IT adoption countries with a group of control observations, drawn from the pool of forecasters in countries that did not adopt IT during the same 24-month period (or the following 12-month period) and which had not adopted IT previously.[17]

The purpose of matching is to eliminate selection bias (due to systematic differences between the treated and untreated observations in the sample). Propensity score matching is a form of matching on observables: the propensity score (estimated probability of being selected into the treatment group) is estimated by running a probit regression on a selection of observable characteristics. Since the observation unit in our dataset is the individual forecaster, forecasters' characteristics are used to estimate the propensity score. This is not to say that the decision to adopt IT is dependent, in a causal sense, on forecaster behavior prior to its adoption. Rather, there may be some third factors (e.g. a crisis prior to the change of monetary policymaking regime) so that prior forecaster behavior differed between countries that adopted or did not adopt IT, biasing the measured impact of IT adoption on subsequent forecast accuracy.

I use eight variables to calculate the propensity score: mean absolute forecast errors for output and inflation in the 12-month period (period 0) up to IT adoption $\left\{V_{0,g}^{ij}; V_0^{ij}\right\}$, the change in these variables from the *previous* 12-month period (period -1) to this period $\left\{\Delta V_{-1,g}^{ij}; \Delta V_{-1}^{ij}\right\}$; and similar variables for the forecast level of these variables

---

[15]The orthogonality assumption can be motivated by invoking the classical dichotomy between real and nominal variables. However, to the extent that forecasters expect some short-term positive relationship between output growth and inflation (i.e., via some kind of Phillips curve), this condition may be violated.

[16]Since shocks to forecast accuracy should, at least in theory, lead forecasters to over- and underestimate inflation with equal probability, f^{ij} should not be correlated with e_0^{ij}.

[17]As it turns out, this aspect of the matching strategy results in no controls drawn from countries that subsequently (in the period covered by our data) adopted IT. This is likely due to the fact that countries that adopted IT generally did so fairly soon after they appeared in the dataset (with advanced countries adopting IT sooner but also appearing in the dataset earlier).

$\{f_{0,g}^{ij}; f_0^{ij}; \Delta f_{-1,g}^{ij}; \Delta f_{-1}^{ij}\}$. Figure 1 shows the distribution of the propensity score within the treatment group and the three control groups (the latter generated according to the methodology outlined in the following paragraphs), confirming that the treatment and control samples are similar in terms of prior characteristics as measured by the propensity score.[18]

Figure 1. Distribution of Propensity Score in Treatment and Control Groups
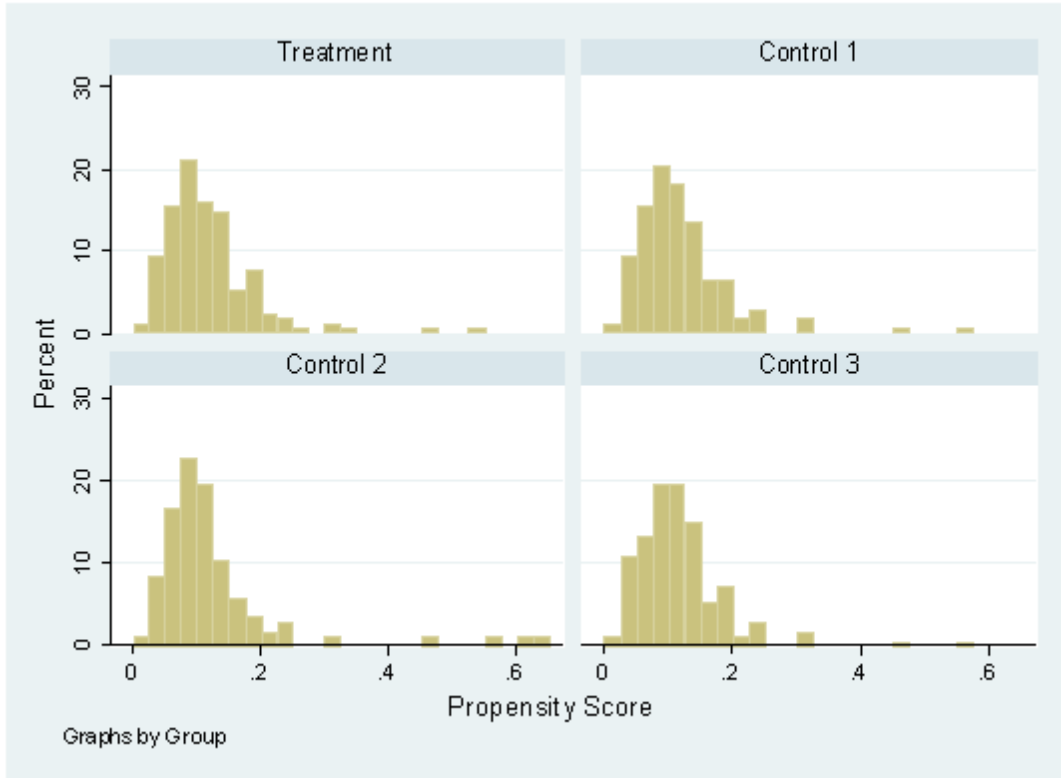
For robustness, I choose three different matching methods. The first chooses the best available control with replacement (so that a single control can appear multiple times as the best match for several of the treated group). This is our preferred method because it identifies the best control for each of the treated observations, regardless of whether the control has already been matched with another observation. The second method is one-to-one matching: treatment observations are matched with the best available match (according to the propensity score) from controls that have not already been matched. One-to-one matching requires that controls are drawn without replacement; it is also particularly important that in this case the data ordering has been randomized. This method has the advantage of simplicity.

---

[18] Observations are frequency weighted (for control groups 1 and 3).

Under these two methods the set of controls for the treated group from a particular country will typically be drawn from several countries. This leads to a very obvious difference between the controls and the treated, since the latter are, for each IT adoption episode, drawn from a single country. This difference could bias the results. Hence, the third method adopts the following two step matching procedure. First, following our preferred method, the best matches are drawn (with replacement) from the full set of available matches. The country providing the highest number of matches (including repeated controls) is then selected for resampling (when countries tie as in one case in our sample—where two countries are tied—all are selected). Controls are then selected from this limited set of observations according to the propensity score (matching with replacement).

Appendix II provides a full account of the matching algorithms used in the paper. Table 1 provides summary statistics (means with standard deviations in parentheses) for the treatment and three control groups. The three samples of treatment and control groups are detailed in Table 2.

Table 1. Descriptive Statistics (frequency weighted)

	Group	Treatment	Control 1	Control 2	Control 3
	V_0^{ij}	2.14	2.93	3.14	3.78
		(1.58)	(4.09)	(4.35)	(5.06)
	ΔV^{ij}	−.934	.237	.0697	−.0465
		(1.45)	(3.15)	(3.54)	(4.04)
Observations:	Unweighted	166	109	166	86
	Weighted	166	166	166	166

C.   Results

The results of estimating equation (13) are presented in Table 3. Panels A-C present results for the three matching methods (1-3 respectively). Model I suppresses the interaction effect, measuring only the levels (or unconditional) effect of IT adoption on forecast errors. Model II estimates the full equation using OLS. Model III uses the IV strategy (two-stage least squares, 2SLS). Since advanced as well as emerging market economies are included in the sample a dummy for advanced countries is included in each specification.[19]

---

[19]For the purposes of this paper, the "advanced" countries in the dataset are Australia, Canada, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Switzerland, the United Kingdom, and the United States.

Table 2. Sample Country Coverage (Unweighted Observations)

PANEL A: Control Group 1

Targeter	Australia	Brazil	Canada	Chile	Colombia	Korea	Mexico	Norway	Peru	Thailand	United Kingdom	Total
Date	Apr-93	Jun-99	Feb-91	Sep-99	Sep-99	Jan-01	Jan-01	Mar-01	Jan-02	May-00	Oct-92	
Total	25	25	31	19	16	24	30	18	20	19	48	275
Treatment	14	14	18	13	10	15	17	10	11	11	33	166
Controls	11	11	13	6	6	9	13	8	9	8	15	109
Argentina		1										1
France	2	2						1	1	1	1	8
Germany	6			1					1	1		9
Hong Kong SAR		1				1		1				3
India		2				3	6	1				12
Indonesia						1		1		1		3
Italy	3										2	5
Japan		1					1					2
Malaysia		2		2	1							5
Netherlands										1		1
Singapore		1							1	1		3
Spain					1		2	2	1			6
Switzerland							1					1
United States		1	13						5	3	12	34
Venezuela				3	4	4	3	2				16

Table 2 (continued). Sample Country Coverage (Unweighted Observations)

PANEL B: Control Group 2

Targeter	Australia	Brazil	Canada	Chile	Colombia	Korea	Mexico	Norway	Peru	Thailand	United Kingdom	Total
Date	Apr-93	Jun-99	Feb-91	Sep-99	Sep-99	Jan-01	Jan-01	Mar-01	Jan-02	May-00	Oct-92	
Total	28	28	36	26	20	30	34	20	22	22	66	332
Treatment	14	14	18	13	10	15	17	10	11	11	33	166
Controls	14	14	18	13	10	15	17	10	11	11	33	166
Argentina		1										1
France	2	2						1	1	2	1	9
Germany	8							1	1	2	1	13
Hong Kong SAR		1				1	1	1				4
India		4				4	7	1		1		17
Indonesia						4	1			1		6
Italy	3										8	11
Japan	1	1		1								3
Malaysia		2		3	2							7
Netherlands										1		1
Singapore		1							1	1		3
Spain					1		2	4	3			10
Switzerland							1					1
United States		1	18						5	3	23	50
Venezuela		1		9	7	6	5	2				30

PANEL C: Control Group 3

Targeter	Australia	Brazil	Canada	Chile	Colombia	Korea	Mexico	Norway	Peru	Thailand	United Kingdom	Total
Date	Apr-93	Jun-99	Feb-91	Sep-99	Sep-99	Jan-01	Jan-01	Mar-01	Jan-02	May-00	Oct-92	
Total	22	21	31	17	16	23	26	15	17	19	45	252
Treatment	14	14	18	13	10	15	17	10	11	11	33	166
Controls	8	7	13	4	6	8	9	5	6	8	12	86
France										2		2
Germany	8											8
India		7					9					16
Spain								5				5
United States			13						6	6	12	37
Venezuela				4	6	8						18

Table 3. Regression Results

ΔV^{ij}	I			II			III		
Model	Levels			Interaction (OLS)			Interaction (2SLS)		
Residuals	iid	robust	clustered	iid	robust	clustered	iid	robust	clustered
Panel A: Nearest Neighbor Matching (with replacement)									
IT	−1.16†††	−1.16†††	−1.16*	.200	.200	.200	.831†	.831†††	.831**
	(.272)	(.305)	(.666)	(.365)	(.245)	(.423)	(.460)	(.234)	(.421)
V_0^{ij}				−.153***	−.153*	−.153	.0382	.0382	.0382
				(.0480)	(.0874)	(.124)	(.0538)	(.118)	(.231)
$IT \times V_0^{ij}$				−.752†††	−.752†††	−.752†††	−.944†††	−.944†††	−.944†††
				(.119)	(.0891)	(.154)	(.180)	(.119)	(.238)
$Adv.$	.0742	.0742	.0742	−.803***	−.803***	−.803	−.306	−.306	−.306
	(.272)	(.316)	(.690)	(.288)	(.292)	(.528)	(.313)	(.239)	(.419)
$Const.$	.193	.193	.193	1.16†††	1.16***	1.16*	.308	.308	.308
	(.250)	(.424)	(.864)	(.319)	(.415)	(.648)	(.347)	(.363)	(.611)
$F - stat$	9.49†††	23.0†††	2.47*	22.6†††	179†††	44.4†††	12.2†††	101†††	22.8†††
R^2	.0545	.0545	.0545	.216	.216	.216	.193	.193	.193
Panel B: One-for-One Nearest Neighbor Matching (without replacement)									
IT	−.961†††	−.961***	−.961	−.0653	−.0653	−.0653	.589	.589**	.589
	(.300)	(.336)	(.729)	(.397)	(.276)	(.489)	(.499)	(.291)	(.462)
V_0^{ij}				−.286†††	−.286†††	−.286**	−.0978*	−.0978	−.0978
				(.0491)	(.0882)	(.112)	(.0547)	(.128)	(.211)
$IT \times V_0^{ij}$				−.637†††	−.637†††	−.637†††	−.820†††	−.820†††	−.820†††
				(.128)	(.0891)	(.152)	(.195)	(.126)	(.231)
$Adv.$	.309	.309	.309	−.973***	−.973***	−.973	−.448	−.448	−.448
	(.300)	(.347)	(.753)	(.313)	(.328)	(.619)	(.339)	(.297)	(.485)
$Const.$	−.112	−.112	−.112	1.54†††	1.54†††	1.54**	.641*	.641	.641
	(.275)	(.473)	(.979)	(.346)	(.470)	(.765)	(.376)	(.461)	(.675)
$F - stat$	6.24***	23.0†††	2.37	23.8†††	170†††	39.7†††	9.17†††	97.4†††	20.0†††
R^2	.0366	.0366	.0366	.225	.225	.225	.210	.210	.210
Panel C: "Best country" Nearest Neighbor Matching (with replacement)									
IT	−.827**	−.827**	−.827	−.389	−.389	−.389	.472	.472*	.472
	(.335)	(.377)	(1.26)	(.440)	(.313)	(.947)	(.550)	(.267)	(.735)
V_0^{ij}				−.365†††	−.365†††	−.365***	−.163*	−.163*	−.163
				(.0477)	(.0631)	(.107)	(.0526)	(.0891)	(.223)
$IT \times V_0^{ij}$				−.598†††	−.598†††	−.598***	−.800†††	−.800†††	−.800***
				(.139)	(.0705)	(.190)	(.215)	(.101)	(.284)
$Adv.$	.453	.453	.453	−1.36†††	−1.36†††	−1.36	−.634	−.634*	−.634
	(.336)	(.388)	(1.30)	(.354)	(.392)	(1.25)	(.387)	(.337)	(1.04)
$Const.$	−.311	−.311	−.311	2.13†††	2.13†††	2.13	.938**	.938**	.938**
	(.306)	(.535)	(1.80)	(.395)	(.548)	(1.66)	(.430)	(.479)	(1.41)
$F - stat$	4.47**	23.6†††	2.22	26.6†††	154†††	26.8†††	8.57†††	97.4†††	16.6†††
R^2	.0265	.0265	.0265	.246	.246	.246	.223	.223	.223

Significance levels in this and subsequent tables are denoted as: ††† 0.1 percent; ***: 1 percent; **: 5 percent; *: 10 percent.

For each model the table presents three sets of results with different standard errors (in parentheses). The first is derived from the standard OLS/2SLS residuals based on iid homoskedastic variance. The second drops the homoskedasticity assumption but maintains independence of the residuals (robust standard errors). The third, our preferred measure, clusters according to IT adoption episode (1-11) and within each episode by country, to control for non-independence of results across forecasters within each country (clustered standard errors).[20]

Bertrand, Dufflo, and Mullainathan (2004) have shown that difference in differences estimation with repeated observations within groups tends to consistently overstate the significance level attached to estimated treatment effects if the within-group correlation is ignored. They show that clustering (e.g., using the "cluster" option in *Stata*, as here) is a simple and effective means of eliminating this problem. The problem is likely to be particularly acute in our data since the observations from the same "group" (defined over episode and country) are *based on forecasts of the same variable*; hence clustering is particularly important in our case.

Table 3 provides extremely strong evidence for the conditional effect predicted by the model (columns II and III), but little evidence for an unconditional effect (column I). The coefficient b_{1T} estimated by OLS and presented in column II is between 0.6 (panel C) and 0.75 (panel A), which is quantitatively significant. Using the clustered standard errors to derive significance levels, the estimates are also highly statistically significant, at least at the 1% level (panel C) and generally at the 0.1% level (panels A and B). The IV results (column III) are even stronger: the point estimate for b_{1T} is between 0.8 and 0.94 and the statistical significance level is the same as with the OLS results. The point estimate derived using 2SLS is higher because the IV strategy succeeds in eliminating the mean reversion present in the OLS results: note that the IV estimate of b_1 is not significantly different from zero whereas the OLS estimate is highly significant as well as having a much higher point estimate.[21] The unconditional effect presented in column I is not generally statistically significant although it carries a negative sign in all specifications.

Table 4 provides some robustness checks, replicating Table 3 for two placebo datasets (presenting clustered standard errors only). Columns I-III replicate Table 3 using data for IT adoption countries with the adoption window shifted to be *12 months earlier than in reality*, and matching controls according to this placebo data.[22] There is no statistically significant evidence for a conditional or unconditional "effect" of the placebo IT adoption variable. Columns IV-VI undertake the same exercise shifting the adoption window *back 12 months*. Again, there is no evidence of an "effect" from the placebo.

Note that since the sample size is greater than that for the genuine data in this latter

---

[20] For the latter, the number of clusters is 58 (panel A), 63 (panel B) and 23 (panel C).

[21] The IV estimates (with clustered residuals) pass tests for instrument relevance (using the Anderson canonical correlations LR statistic) and exogeneity (Hansen's J statistic used for the Hansen-Sargan overidentification test). They also pass Stock and Yogo's tests for weak instruments based on 2SLS bias and size (although these tests strictly require homoskedastic residuals; see Stock and Yogo, 2002; Stock, Wright, and Yogo, 2002).

[22] Note that the first adoption episode (Canada) drops out of the sample due to data constraints.

placebo experiment, but lower in the first placebo experiment, differences in data coverage seem unlikely to explain the difference between the genuine and the placebo regressions. Table 4 therefore provides strong evidence that the results in Table 3 are not due to chance or to country-specific factors relating to the treatment (IT adoption) group, but rather genuinely capture the conditional effect of IT adoption on forecast errors predicted by the model.

Table 5 provides additional robustness checks. Despite the use of propensity score matching to ensure broadly similar characteristics in treatment and control observations, the average prior forecast error is somewhat higher in the control group due to a number of particularly high observations (including some that are above the maximum value for this variable in the treatment group). There is no a priori reason why this would bias the results toward finding a (spurious) conditional effect of IT, particularly since greater variation in the control data might have been expected to lead to greater (or at least better estimated) mean reversion for these observations and hence bias the results *away* from finding greater covergence in the treatment group. However, to ensure that this is not driving the results in Table 3, Table 5 presents results for three trimmed samples in which some controls with high values for V_0^{ij} have been dropped (again, clustered standard errors are presented).

In columns I-III the 5 percent of observations with the highest values for V_0^{ij} (all controls) are dropped from the dataset. In columns IV-VI all controls with values for V_0^{ij} above the maximum value for the treatment group (7.4) are dropped. Eliminating these controls has little substantial effect on the measured conditional effect of IT. The point estimate for b_{1T} is somewhat higher, if anything, but the significance level is broadly unchanged except for the OLS estimates (columns II and V) in panel C, where the estimated effect is no longer statistically significant. On the other hand, the IV estimates of b_{1T} (columns III and VI) are strengthened in terms of both point estimates and statistical significance.[23] Finally, columns VII-IX present results with control observations from Venezuela dropped. This country accounts for the majority of outliers, both in terms of V_0^{ij} and the dependent variable, ΔV^{ij}. Dropping these observations reduces the point estimate on b_{1T} but the coefficient remains significantly different from zero except for that presented in panel C (columns VIII and IX).[24]

Overall, the results offer strong support for the conditional effect predicted by the model. There is little or no evidence of an unconditional effect, confirming the results of Johnson (2002).

---

[23] Applying IV to these restricted samples succeeds in not only eliminating mean reversion, but finds strong evidence for the opposite phenomenon — bad forecasters deliver worse forecasts subsequently ($-b_1 > 0$) — with the adoption of IT restoring mean reversion ($-(b_1 + b_{1T}) < 0$). This effect, if genuine, merits further study.

[24] Even in this case the p–value associated with the IV estimates (.122) indicates borderline statistical significance.

Table 4. Placebo Regression Results

ΔV^{ij}	I	II	III	IV	V	VI
Placebo		12 Months Before			12 Months After	
Model	Levels	OLS	2SLS	Levels	OLS	2SLS
Panel A: Nearest Neighbor Matching (with replacement)						
IT	−.648	.617	.0828	−.115	.218	.705
	(.495)	(.556)	(.868)	(.395)	(.455)	(.736)
V_0^{ij}		.152	.360		.0312	.375
		(.235)	(.246)		(.187)	(.241)
$IT \times V_0^{ij}$		−.457	−.209		−.273	−.588
		(.284)	(.313)		(.494)	(.674)
$Adv.$	−.322	−.227	.471	−.219	−.194	.0715
	(.423)	(.419)	(.557)	(.403)	(.442)	(.490)
$Const.$	.308	−.163	−1.08	.296	.238	−.390
	(.505)	(.701)	(.768)	(.380)	(.475)	(.603)
$F-stat$	2.33*	1.44	.64	.20	.42	.86
R^2	.0206	.0752	< 0	.0057	.0114	< 0
Clusters		57			69	
Panel B: One-for-One Nearest Neighbor Matching (without replacement)						
IT	−1.05	−.180	−.0506	−.340	−.0737	.808
	(.652)	(.646)	(.830)	(.435)	(.479)	(.765)
V_0^{ij}		−.0441	.254*		−.0129	.494
		(.140)	(.152)		(.222)	(.349)
$IT \times V_0^{ij}$		−.352	−.251		−.229	−.820
		(.239)	(.260)		(.503)	(.706)
$Adv.$	−.638	−1.01*	−.00165	−.470	−.481	−.0656
	(.544)	(.592)	(.596)	(.450)	(.457)	(.514)
$Const.$	.812	1.13	−.400	.623	.647	−.305
	(.708)	(.834)	(.762)	(.472)	(.482)	(.706)
$F-stat$	1.35	1.37	.99	.76	.87	.87
R^2	.0329	.0514	< 0	.0224	.0264	< 0
Clusters		62			73	
Panel C: "Best country" Nearest Neighbor Matching (with replacement)						
IT	−.171	−.681	.434	.524	.178	.300
	(.412)	(.492)	(1.18)	(.415)	(.532)	(.760)
V_0^{ij}		−.737†††	.554		−.395	−.326
		(.111)	(.810)		(.284)	(.312)
$IT \times V_0^{ij}$		.356	−.352		.153	.0742
		(.237)	(.626)		(.577)	(.699)
$Adv.$	.111	−.874**	.798	.294	−.00572	.0468
	(.329)	(.399)	(1.10)	(.391)	(.462)	(.515)
$Const.$	−.306	1.54***	−1.67	−.554*	.202	.0698
	(.342)	(.497)	(1.97)	(.296)	(.576)	(.681)
$F-stat$	.12	17.5†††	.13	1.41	2.04	1.23
R^2	.0063	.225	< 0	.0649	.131	.146
Clusters		20			26	
Obs. (weighted)	272	272	272	352	352	352

IV. Conclusions

The issue of transparency has become central to discussions of central bank governance in academia and among policymakers. The consensus view of IT as a monetary policy framework is that it delivers enhanced transparency as a significant benefit. However, this is ultimately an empirical question, and little empirical work has, up to now, been conducted on the issue.

This paper outlines a simple signal-extraction model for analyzing these issues and derives a testable proposition: if IT enhances transparency in the manner assumed in the model,

then its introduction should promote convergence to lower forecast errors. In other words, forecast errors should decline under IT, *proportionately to the forecasters' initial errors.* This conditional result, derived from a micro-founded model of rational forecasters, is the correct prediction to take to the data.

I test this proposition using matched difference-in-differences, identifying a window around the adoption of IT in 11 countries and matching forecasters in these countries with their counterparts in similar countries that did not adopt IT via three different propensity score matching strategies. I find that convergence occurs in all countries, due to mean-reversion, but that the adoption of IT leads to greater convergence, as predicted by the model. This effect is largely robust to dropping subsets of controls. Moreover, the effect is absent when placebo regressions (with the timing of IT's adoption shifted by a year before or after the genuine date) are run. Finally, I am able to eliminate (non-IT-specific) mean reversion by instrumenting for the initial forecast accuracy. However, the estimated conditional effect of IT adoption is not eliminated by adopting this IV strategy: if anything, it is strengthened. I interpret these results as strong evidence that IT does indeed enhance transparency.

The significant conditional effect of IT ($-b_{1T} < 0$) is in the spirit of Morris and Shin's (2002) argument that better public information is most beneficial for private forecasters whose own information is bad. However, the levels effect b_{0T} is not significantly different from zero (when the interaction effect is included). Hence, there is no evidence that IT adoption can lead to *higher* forecast errors, even for the very best forecasters with initial forecast errors aready close to zero. Assuming that IT increases transparency, this finding goes against Morris and Shin's argument that better public information can make private forecasts less accurate, or at least suggests that the rather special conditions (e.g. restrictions on parameter values) necessary for this case to hold are absent for the forecasters in our sample.

Further research could use the same dataset and techniques to test whether IT enhances transparency with respect to other variables. Preliminary results for output growth suggest that IT adoption has little or no effect on forecaster behavior, perhaps unsurprisingly since the monetary authorities' control and forecasting advantage over real variables is likely far lower than for inflation. The results for inflation are perhaps most interesting in any case, since protecting the real value of money is typically the central bank's primary objective. Moreover, since inflation expectations play such a critical role in the monetary transmission mechanism, establishing some transparency benefits of IT in this area is probably of greatest interest to policymakers. In this regard, the results presented here constitute fairly compelling evidence for the transparency benefits of IT.

Table 5. Robustness Checks

ΔV^{ij}	I	II	III	IV	V	VI	VII	VIII	IX
Criterion	$F(V_0^{ij}) \le .95$			$V_0^{ij} \le \max(V_0^{ij} \mid IT=1)$			Venezuela dropped		
Model	Levels	OLS	2SLS	Levels	OLS	2SLS	Levels	OLS	2SLS
Panel A: Nearest Neighbor Matching (with replacement)									
IT	−1.53** (.696)	1.35††† (.371)	3.04††† (.571)	−1.26** (.518)	.992*** (.456)	3.30††† (.883)	−.529 (.373)	.241 (.224)	.746** (.322)
V_0^{ij}		.753* (.416)	1.94††† (.272)		.395 (.467)	2.17††† (.677)		−.560††† (.0757)	−.285 (.250)
$IT \times V_0^{ij}$		−1.59††† (.424)	−2.66††† (.265)		−1.25*** (.466)	−2.88††† (.638)		−.290††† (.0811)	−.594** (.246)
$Adv.$	−.315 (.751)	−.186 (.311)	.896 (.571)	−.0332 (.585)	−.294 (.327)	1.01 (.663)	.729 (.452)	−.274 (.243)	−.0579 (.255)
$Const.$	.744 (.910)	−.407 (.439)	−2.84††† (.843)	.345 (.651)	.0275 (.558)	−3.17*** (1.23)	−.735 (.471)	.765** (.304)	.226 (.424)
$F-stat$	2.98*	84.1†††	42.8†††	3.32**	75.1†††	28.7†††	2.41*	91.7†††	29.2†††
R^2	.106	.415	.156	.0920	.350	< 0	.138	.771	.769
Obs. (Clusters)		307 (54)			304 (54)			305 (53)	
Panel B: One-for-One Nearest Neighbor Matching (without replacement)									
IT	−1.30* (.711)	.644* (.363)	2.35** (.924)	−1.31** (.630)	.836* (.464)	3.04††† (.828)	−.390 (.376)	.0618 (.217)	.684** (.337)
V_0^{ij}		.0858 (.352)	.948 (.577)		.236 (.444)	1.57*** (.558)		−.707††† (.0811)	−.377 (.264)
$IT \times V_0^{ij}$		−.950** (.363)	−1.68††† (.522)		−1.09** (.444)	−2.27††† (.526)		−.144* (.0849)	−.500* (.257)
$Adv.$	−.0398 (.750)	−.407 (.337)	.952 (.767)	−.0551 (.681)	−.314 (.335)	1.22* (.652)	.898* (.459)	−.283 (.240)	.00261 (.251)
$Const.$	.382 (.943)	.451 (.436)	−2.15 (1.38)	.404 (.823)	.197 (.565)	−3.03*** (1.14)	−.949* (.486)	.950*** (.298)	.255 (.443)
$F-stat$	2.62*	66.8†††	30.2†††	2.89*	73.6†††	25.2†††	2.59*	98.5†††	31.0†††
R^2	.0658	.217	.0085	.0804	.279	< 0	.138	.766	.759
Obs. (Clusters)		316 (54)			308 (61)			302 (57)	
Panel C: "Best country" Nearest Neighbor Matching (with replacement)									
IT	−1.53 (1.42)	.561 (.582)	3.02** (1.20)	−1.41 (1.16)	.570 (.626)	2.92*** (.998)	−.106 (.773)	.106 (.312)	.470 (.248)
V_0^{ij}		.148 (.760)	1.41** (.589)		.137 (.820)	1.56** (.612)		−.757††† (.0891)	−.600††† (.223)
$IT \times V_0^{ij}$		−1.03 (.782)	−2.07††† (.618)		−1.02 (.839)	−2.27††† (.641)		−.0859 (.0616)	−.279 (.180)
$Adv.$	−.274 (1.48)	−.579 (.457)	1.51 (1.56)	−.143 (1.22)	−.555 (.434)	1.15 (1.25)	1.19 (.846)	−.211 (.362)	−.0592 (.295)
$Const.$	.724 (2.03)	.651 (.607)	−.323 (2.20)	.537 (1.65)	.626 (.639)	−2.85 (1.77)	−1.37 (1.11)	.858 (.499)	.502 (.343)
$F-stat$	1.92	53.0†††	20.1†††	2.07	54.6†††	24.8†††	2.58	173†††	30.9†††
R^2	.0710	.210	< 0	.0720	.236	< 0	.152	.803	.824
Obs. (Clusters)		306 (21)			298 (21)			294 (20)	

APPENDIX I.   COMPARATIVE STATICS IN THE MODEL OF MORRIS AND SHIN (2002)

Morris and Shin's (2002) model of strategic behavior in response to public and private information is well suited to the task of modeling the private sector's forecasts given (a) competing public (central bank) and private information and (b) potential benefits from matching others' forecasts (the authors themselves have this specific example in mind, although the model applies more widely). The model's strategic component is a formalization of Keynes's (1936) famous "beauty contest" example, which Keynes proposed as a metaphor for financial market behavior. In Section II we present the model with strategic behavior eliminated—leading to a simpler model of optimal signal extraction. Here we present the model with the strategic component included to illustrate under what conditions the results discussed in Section II hold for the more complex model. The brief exposition below is simply a restatement of the model in its original setting, with the terminology modified to more explicitly match our narrower focus on forecasting inflation.[25]

A continuum of agents ("forecasters") are distributed on the unit interval. Their preferences from forecasting inflation π depend on the accuracy of their forecast f_i and the relative closeness of their forecast with respect to others' forecasts (the beauty contest element).[26] Specifying the action profile (i.e., the vector of forecasts) over all agents as $\mathbf{f}$, the preferences of forecaster i can be given as:

$$u_i\left(\mathbf{f}, \pi\right) \equiv (1 - r)\left(f_i - \pi\right)^2 - r\left(L_i - \overline{L}\right), \qquad (14)$$

where $r \in [0, 1]$, a constant, denotes the relative weight placed on the "beauty contest" element and L_i denotes the loss from differing from others' forecasts:

$$
\begin{aligned}
L_i &\equiv \int_0^1 \left(f_j - f_i\right)^2 dj \\
\overline{L} &\equiv \int_0^1 L_j dj.
\end{aligned}
$$

The private sector agents observe the central bank's public signal (a combination of its public forecasts, statements and analysis), π_C, and also observe their own private signal π_i. Each signal is noisy with precision (the inverse of the variance) denoted, respectively, as α and β as in Section II.

Morris and Shin demonstrate that in the unique equilibrium agents weight the two signals according to their relative precision and the importance of the beauty contest element

---

[25] Refer to Morris and Shin (2002) and Morris, Shin and Tong (2006) for proofs and further details.

[26] As Morris and Shin (2002) note, the beauty contest motivation in forecasters' behavior imposes an externality: good forecasters of the average forecast gain at the expense of bad forecasters of the average forecast, but overall welfare is unchanged by definition. In fact normalized overall welfare is simply (minus) the mean squared forecast error: $\frac{W(\mathbf{f}, \pi)}{1 - r} = -\int_0^1 \left(f_i - \pi\right)^2 di$.

compared with the importance of arriving at an accurate forecast:[27]

$$f_i = \frac{\alpha \pi_C + \beta (1 - r) \pi_i}{\alpha + \beta (1 - r)}. \tag{15}$$

Then the expected mean square forecast error (a measure of the forecast inaccuracy) is given by:

$$E\left[(f_i - \pi)^2\right] = \frac{\alpha + \beta (1 - r)^2}{(\alpha + \beta (1 - r))^2}. \tag{16}$$

The private forecasts become more accurate (forecast error falls) as the public signals become more precise, if the beauty contest element is not too prominent ($r \leq 0.5$) or, even if $r > 0.5$, if the public information is relatively precise compared to the private information:

$$\frac{\partial E\left[(f_i - \pi)^2\right]}{\partial \alpha} = -\frac{\alpha - (2r - 1)(1 - r)\beta}{(\alpha + \beta (1 - r))^3} \tag{17}$$

$$\lesseqgtr 0 \text{ as } \frac{\alpha}{\beta} \gtreqless (2r - 1)(1 - r).$$

Hence:

Prediction (i) For the worst forecasters ($\beta < 8\alpha$), better public information always leads to reduced forecast errors;

Prediction (ii) When the beauty contest element is important ($r \geq 0.5$), then enhanced transparency can lead to higher forecast errors for the best forecasters ($\beta > \frac{\alpha}{(2r-1)(1-r)}$).

More generally, differentiating the preceeding expression with respect to β yields the following:

$$\frac{\partial^2 E\left[(f_i - \pi)^2\right]}{\partial \alpha \partial \beta} = \frac{2(1 - r)(\alpha (1 + r) - (2r - 1)(1 - r)\beta)}{(\alpha + \beta (1 - r))^4} \tag{18}$$

$$\gtreqless 0 \text{ as } \frac{\alpha}{\beta} \gtreqless \frac{(2r - 1)(1 - r)}{(1 + r)}.$$

The effect of transparency on forecast errors increases (i.e. *falls in absolute terms* since the effect is negative for low values of β) as the accuracy of the private signals β increases,

---

[27] Greater weight is put on the public signal when the beauty contest motivation is more prominent: this is because the central bank's signal serves as a focal point for higher order beliefs (forecasts of forecasts of forecasts etc.). Essentially, when trying to second guess other forecasters, the private sector places extra weight on the public signal because it knows that it is directly observed by the other forecasters.

again if the beauty contest element is not too prominent ($r \leq 0.5$) or if the public information is relatively precise compared to the private information. This corresponds to the hypothesis $-b_1 < 0$ in Section II.

As Svensson (2006) has noted, realistic parameter values would imply that $\frac{\alpha}{\beta} > (2r - 1)(1 - r)$.[28] This would lead us to expect prediction (ii) to not hold in the data.

---

[28] As Svensson (2006) notes, $\frac{\alpha}{\beta} < (2r - 1)(1 - r)$ would require $\frac{\alpha}{\beta} < \frac{1}{8}$ even for the value of r (0.75) that maximizes $(2r - 1)(1 - r)$; it seems unlikely that the private sector's signal would be on average more than eight times as accurate as that of the Central Bank (particularly since Romer and Romer's (2000) evidence on the relative accuracy of official and private sector forecasts in the US suggests that the opposite scenario, in terms of relative forecast accuracy, is more likely). The condition $\frac{\alpha}{\beta} < \frac{(2r-1)(1-r)}{(1+r)}$ for $\frac{\partial^2 E\left[(f_i - \pi)^2\right]}{\partial \alpha \partial \beta} > 0$ to *not* hold is even more restrictive.

Appendix II. Matching Algorithms

For a general discussion of propensity score matching, including a comparison of matching with and without replacement, see Smith and Todd (2005). For all three matching algorithms, the first stage is to estimate the propensity score (ps). The relevant sample is drawn, for the 11 IT adoption episodes for which we have data, from (a) the country that adopted IT in the particular episode (the treatment observations); and (b) all other countries in the dataset that did not adopt IT during the 24-month window defining the episode or the subsequent 12-month period and had not adopted IT prior to this episode (the pool of potential controls). ps is then estimated by running a probit regression with the eight right-hand-side variables described in Section III B $\left\{ V_{0,g}^{ij}, V_{0}^{ij}, \Delta V_{-1,g}^{ij}, \Delta V_{-1}^{ij}, f_{0,g}^{ij}, f_{0}^{ij}, \Delta f_{-1,g}^{ij}, \Delta f_{-1}^{ij} \right\}$ and taking the fitted probability. The total number of observations used to estimate ps is 2,141.[29] Matching is undertaken in *Stata* using the *psmatch2* command (Leuven and Sianesi, 2003). In all cases (except the second step in the third algorithm) a common support in terms of ps is required (treatment observations with a ps value outside the support of ps in the control group are dropped).

Algorithm 1: nearest-neighbor matching (with replacement)

Observations are ordered randomly. For each treatment observation, the nearest neighbor (least absolute distance in terms of ps) *drawn from within the same episode group* is chosen as the control observation (if there is a tie, the first available observation is chosen). Matching is undertaken with replacement, so that some controls appear as repeated observations in the dataset (weighted according to frequency to give an effective dataset size of 332).

Algorithm 2: one-to-one matching (without replacement)

As with Algorithm 1, except that matching is now undertaken without replacement, so that there is a unique correspondence between the 166 control and 166 treatment observations (once a match has been made, the control observation is removed from the pool of potential controls before the match for the next treatment observation is sought).

Algorithm 3: nearest-neighbor matching (with replacement) from the "best" available country or countries only

Step 1 replicates algorithm 1.

Step 2: pick the "best" country as that with the highest number of (frequency weighted) matches in step 1. If there is a tie (as for one episode in our data) pick both countries. Now repeat the matching algorithm (again, matching with replacement according to ps), *using only forecasters from these best countries as the pool of potential controls.*

---

[29]The probit regression has a $\chi^2(8)$ statistic of 87.6 [p-value=0.000] and a pseudo-R^2 of .07. As one might expect, the four right-hand-side variables associated with forecasts of inflation are more correlated with the IT adoption decision (reflected in higher z-statistics) than their counterparts derived from output growth forecasts.

REFERENCES

Ang, Andrew, Geert Bakaert, and Min Wai, 2005, "Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?" NBER Working Paper No. 11538 (Cambridge, Massachusetts: National Bureau of Economic Research).

Ball, Laurence, and Niamh Sheridan, 2004, "Does Inflation Targeting Matter?" in *The Inflation Targeting Debate*, ed. by Bernanke and Woodford (Chicago: University of Chicago Press).

Barro, Robert, and David Gordon, 1983, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy*, Vol. 91, No. 4, pp. 589–610.

Berg, Claes, 2005, "Experience of Inflation-Targeting in 20 Countries," *Sveriges Riksbank Economic Review*, Vol. 1, pp. 20–47.

Berger, Helge, Michael Ehrmann, and Marcel Fratzscher, 2006, "Forecasting ECB Monetary Policy: Accuracy Is (Still) a Matter of Geography," IMF Working Paper 06/41 (Washington: International Monetary Fund).

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics,* Vol. 119, pp. 249–75.

Besley, Timothy, and Anne Case, 2000, "Unnatural Experiments? Estimating the Incidence of Endogenous Policies," *Economic Journal,* Vol. 110, pp. F672–94.

Bernanke, Ben, Thomas Laubach, Frederic Mishkin, and Adam Posen, 1999, *Inflation Targeting: Lessons from the International Experience* (Princeton: Princeton University Press).

Blinder, Alan, Charles Goodhart, Philipp Hildebrand, David Lipton, and Charles Wyplosz, 2001, *How Do Central Banks Talk*? Geneva Report on the World Economy 3 (London: Centre for Economic Policy Research).

Carpenter, Seth, 2004, "Transparency and Monetary Policy: What Does the Academic Literature Tell Policymakers" (unpublished; Board of Governors of the Federal Reserve System).

Chortareas, Georgios, David Stasavage, and Gabriel Sterne, 2002, "Does It Pay to be Transparent? International Evidence from Central Bank Forecasts," *Federal Reserve Bank of St. Louis Review,* July/August 2002, pp. 99–118.

Corbo, Vittorio, Oscar Landerretche, Klaus Schmidt-Hebbel (2001), "Assessing Inflation Targeting after a Decade of World Experience," *International Journal of Finance and Economics* 6 (4), pp. 343-368.

Cukierman, Alex, and Allan Meltzer, 1986, "A Theory of Ambiguity, Credibility, and Inflation Under Discretion and Asymmetric Information," *Econometrica*, Vol. 54, No. 5, pp. 1099–1128.

Eijffinger, Sylveser, and Petra Geraats, 2006, "How Transparent are Central Banks?" *European Journal of Political Economy,* Vol. 22, No. 1, pp. 1–21.

Faust, Jon, and Dale Henderson, 2004, "Is Inflation Targeting Best-Practice Monetary Policy?" *Federal Reserve Bank of St. Louis Review*, Vol. 86, pp. 117–44.

Fry, Maxwell, Deanne Julius, Lavan Mahadeva, Sandra Roger, and Gabriel Sterne, 2000, *Key Issues in the Choice of Monetary Policy Framework* (Mahadeva and Sterne).

Geraats, Petra, 2002, "Central Bank Transparency," *Economica Journal*, Vol. 112, pp. F532–65.

Goodhart, Charles, and Ellen Meade, 2004, "Central Banks and Supreme Courts," *Moneda y Credito*, Vol. 218 (Madrid: Fundacion Santander Central Hispano).

Gurkaynak, Refet, Andrew Levin, and Eric Swanson, 2005, "Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from Long-Term Bond Yields in the U.S., U.K., and Sweden." Mimeo.

Johnson, David, 2002, "The Effect of Inflation Targeting on the Behavior of Expected Inflation: Evidence from an 11 Country Panel," *Journal of Monetary Economics,* Vol. 49, pp. 1521–38.

Keynes, John Maynard, 1936, *The General Theory of Employment, Interest and Money* (London: Macmillan).

King, Mervyn, 1997, The Inflation Target Five Years On (lecture delivered at the LSE, October 29 (unpublished; London, Bank of England).

Kuttner, Kenneth, and Adam Posen, 1999, "Does Talk Matter After All? Inflation Targeting and Central Bank Behavior," Staff Report 88, Federal Reserve Bank of New York.

Kydland, Finn, and Edward Prescott, 1977, "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy,* Vol. 85, No. 3, pp. 473–91.

Leuven, Edwin, and Barbara Sianesi, 2003, "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing." Available via Internet: http://ideas.repec.org/c.boc/bocode/s432001.html. Version 3.0.0.

Mahadeva, Lavan, and Gabriel Sterne, 2000, *Monetary Policy Frameworks in a Global Context* (London: Routledge).

Meade, Ellen, 2006, "Dissents and Disagreement on the Fed's FOMC: Understanding Regional Affiliations and Limits to Transparency," Netherlands Central Bank Research Department Working Paper 094.

Mishkin, Frederic, and Klaus Schmidt-Hebbel, 2001, "One Decade of Inflation Targeting in the World: What Do We Know and What Do We Need to Know?" NBER Working Paper No. 8397 (Cambridge, Massachusetts: National Bureau of Economic Research).

Morris, Stephen, and Hyun Song Shin, 2002, "Social Value of Public Information," *American Economic Review*, Vol. 92, No. 5, pp. 1521–34.

———, and Hui Tong, 2006, "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply," *American Economic Review,* Vol. 96, No. 1, pp. 453–55.

Petursson, Thorarinn, 2004, "The Effects of Inflation Targeting on Macroeconomic Performance," Central Bank of Iceland Working Paper No. 23.

Roger, Scott, and Mark Stone, 2005, "On Target? The International Experience with Achieving Inflation Targets," IMF Working Paper 05/163 (Washington: International Monetary Fund).

Romer, Christine, and David Romer, 2000, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, Vol. 90, No. 3, pp. 429–57.

Smith, Jeffrey A., and Petra E. Todd, 2005, "Does Matching Address Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*, Vol. 125, No. 2, pp. 305–53.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo, 2002, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, Vol. 20, No. 4, pp. 518–29.

Stock, James H., and Motohiro Yogo, 2002, "Testing for Weak Instruments in Linear IV Regression," NBER Technical Working Paper No. 284 (Cambridge, Massachusetts: National Bureau of Economic Research).

Svensson, Lars, 1999, "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics*, Vol. 43, pp. 607–54.

———, "Inflation Targeting: Should it be Modeled as an Instrument Rule or a Targeting Rule?" NBER Working Paper No. 8925 (Cambridge, Massachusetts: National Bureau of Economic Research).

———, "Social Value of Public Information: Comment: Morris and Shin (2002) "Is Actually Pro-Transparency, not Con," *American Economic Review*, Vol. 96, No. 1, pp. 448–52.

Vega, Marco, and Diego Winkelried, 2005, "Inflation Targeting and Inflation Behavior: A Successful Story?" *International Journal of Central Banking,* Vol. 1, No. 3, pp. 153–75.