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

Christopher Crowe[†]

Research Department, International Monetary Fund

700 19th Street NW, Washington, DC 20431.

ccrowe@imf.org.

November 19, 2008

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.

[†]This work reflects the views of the author alone and does not reflect the views of the IMF, its Executive Board or Management. The author would like to thank, subject to the usual caveats, Paolo Mauro, Ellen Meade, Scott Roger, Philip Schellekens, Hyun Shin and participants at seminars at American University, the IMF and the Bank of England and the Royal Economic Society's 2008 Annual Conference for comments on earlier drafts of the paper, and Martin Minnoni for excellent research assistance.

JEL Classifications: D82, E52, E58, G14.

Keywords: Inflation Targeting, Central Bank Transparency, Inflation Forecasts, Propensity Score Matching.

I. Introduction

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

The consensus view among both policymakers and academics is that the introduction of inflation targeting (IT) increases the transparency of monetary policymaking (Bernanke and others, 1999; Faust and Henderson, 2004; King, 1997; Mishkin and Schmidt-Hebbel, 2001; Svensson, 1999). Following Geraats (2002), transparency can be thought of as "the removal of information asymmetries." Geraats catalogues five areas of monetary policymaking where transparency could effect outcomes: (1) the central bank's objectives and its institutional relationship with the rest of government, (2) the publication of data and forecasts, (3) internal decisionmaking, (4) communication and explanation of policy changes and (5) 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).[1]

Ultimately, the question of whether IT enhances Central Bank transparency, and with what effects, is an empirical one. Several papers have attempted to measure the transparency-enhancing effects of IT by focusing on economic outcomes.[2] Gurkaynak, Levin, and Swanson (2005) find that IT appears to make 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) 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).

Other papers use subjective assessments of transparency practices as the benchmark for assessing the impact of IT. Eijffinger and Geraats (2006) analyze the practices of nine

[1] The theoretical literature on the potential benefits of enhanced central bank transparency has, over time, shifted in favor of more transparency, as the potential upside from opacity (the ability to deliver monetary surprises) has been downplayed and the long-term credibility gains of more transparent policymaking have become more salient (Geraats, 2002 and Carpenter, 2004). However, Morris and Shin (2002) dissent from the emerging consensus, arguing that 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. Svensson (2006) shows that this result relies on unlikely parameter values; in the more general case the central bank can publish and not be damned. Ottavianni and Sørensen (2006) provide a further discussion of the impact of strategic forecasting behavior.

[2] The papers cited here focus on transparency-related questions. Others have investigated the behavior of inflation under IT more generally (Ball and Sheridan, 2004; Kuttner and Posen, 1999; Lin and Ye, 2007; Petursson, 2004; Vega and Winkelried, 2005) with mixed results.

major central banks and conclude 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. Crowe and Meade (2007) find that transparency increased among central banks that introduced IT. However, 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. Meanwhile, other work has attempted to match subjective transparency assessments to outcomes. Chortareas, Stasavage, and Sterne (2002, 2003) find that greater transparency, in the form of more public prominence for the central bank's forecasts, is associated with lower inflation and lower unemployment costs of disinflation in a diverse cross-section of countries.[3] Crowe and Meade (2008) find that greater central bank transparency is associated with private sector forecasters making more use of public information. Swanson (2004) finds that financial market forecasts of policy interest rates in the United States improved markedly over time, in step with transparency-enhancing changes to the Fed's communications strategy and policymaking.

This paper offers a direct test of the transparency benefits of IT. It uses the Consensus Economics dataset, focusing on medium-term inflation forecasts by individual forecasters. The use of individual forecasters (rather than averages per country) allows one to test whether the effect of IT differs across forecasters with different characteristics. This is useful, since the simple signal extraction model outlined in section II predicts that IT's impact on forecast accuracy will be conditional on how accurate the forecasts are in the first place. Using individual forecasters as the unit of observation allows us to estimate this conditional effect. A critical issue in assessing the effect of IT is that the assignment of the

[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).

'treatment' (in this case, IT adoption in the country in question) is likely to be non-random. To control for endogenous IT adoption, forecasters in IT adoption countries are matched with a control group of forecasters in non-IT adoption countries using propensity score matching. A number of papers in the literature on IT have started to adopt these techniques for dealing with non-random adoption of IT (Lin and Ye, 2007, Diego and Winkelried, 2005). However, this paper is the first to apply these techniques to the question of IT's transparency benefits. It is also the first to use forecaster-level data and therefore uses forecaster characteristics to undertake the match. A second econometric problem (endogeneity resulting from the inclusion of the initial forecast error, which leads to mean-reversion) is dealt with using IV estimation. The effect of IT adoption is then estimated by comparing the behavior of forecasters' errors in the twelve months leading up to and following the adoption of IT in each country. The sample includes 166 forecasters across 11 IT-adoption episodes, with up to 166 forecasters from non-IT adoption countries forming the control group.

The results are strongly supportive of IT adoption leading to better private sector forecasts. Moreover, the effect is strongest for forecasters whose initial forecast accuracy is worst, in line with the model. This finding is subjected to several robustness checks, and found to be robust. However, while IT adoption improves the forecast accuracy most for the worst forecasters, there is no evidence that the best forecasters are harmed. Hence, the results are not supportive of Morris and Shin's (2002) concerns over transparency's potential downsides.[4]

The rest of the paper is organized as follows. Section II outlines a simple signal extraction model with public and private information to motivate the discussion and provide some predictions. Section III desribes the data used. Section IV outlines the methodology, while section V gives the results and section VI concludes.

[4]See footnote 1.

1 II. Theoretical framework

I motivate the empirical analysis via a simple signal extraction model. 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):

$$u_i\left(f_i, \pi\right) \equiv -\left(f_i - \pi\right)^2. \tag{1}$$

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:

$$\begin{aligned} \pi_C &= \pi + \eta \\ \pi_i &= \pi + \varepsilon \end{aligned} \tag{2}$$

and the precision of the public and private signals is denoted, respectively, as:[5]

$$\begin{aligned} \alpha &= \frac{1}{\sigma_\eta^2} \\ \beta &= \frac{1}{\sigma_\varepsilon^2}. \end{aligned} \tag{3}$$

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

$$f_i^* = \frac{\alpha \pi_C + \beta \pi_i}{\alpha + \beta}. \tag{4}$$

[5] Romer and Romer (2000) analyze whether central bank forecasts (specifically, the Federal Reserve's forecasts that are published only with a 5 year lag) are actually superior to the professional forecasts of the private sector, and they find persuasive evidence that this is the case ($\alpha > \beta$). 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. 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.

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

$$\widetilde{V} \equiv E\left[(f_i - \pi)^2\right] = \frac{1}{(\alpha + \beta)}. \tag{5}$$

In order to take this relation to the data, I introduce some identifying assumptions: (a) that the precision of the private signals is constant, for each forecaster i, over the two year time period covered in the empirical section; and (b) and that the precision of the public signal depends, over the same two year period, only on some country-specific factor and whether the central bank of country j has adopted inflation targeting ($IT = \{0, 1\}$):

$$\begin{aligned} \alpha_t^j &= \alpha^j(IT) \\ \beta_t^i &= \beta^i \end{aligned} \tag{6}$$

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:

$$\left(\widetilde{V}^{ij} \mid IT = 0\right) \equiv \widetilde{V}_0^{ij} = \frac{1}{\left(\alpha^j(0) + \beta^i\right)} \tag{7}$$

and hence:

$$\frac{\partial}{\partial \alpha^j}\widetilde{V}^{ij} = -\frac{1}{\left(\alpha^j(0) + \beta\right)^2} = -\left(\widetilde{V}_0^{ij}\right)^2 < 0 \tag{8}$$

$$\frac{\partial^2 \widetilde{V} ij}{\partial \alpha^j \partial \widetilde{V}_0^{ij}} = -2\widetilde{V}_0^{ij} < 0 \tag{9}$$

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

$$\begin{aligned} \Delta\widetilde{V}^{ij} &\equiv \widetilde{V}_1^{ij} - \widetilde{V}_0^{ij} \simeq b_0 - b_1\widetilde{V}_0^{ij} \\ -b_1 &< 0 \end{aligned} \tag{10}$$

III. Data

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).[6] 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.[7]

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).[8] For each country there is a panel of forecasters whose composition changes

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

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

[8] 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.

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.[9] 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.[10] 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. The 166 forecasters for whom we have "before" and "after" data in the window around IT's adoption, in the 11 countries that adopt IT, form the "treatment" group.

IV. Methodology

A. Specification

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 (10) is given by:

$$\Delta V^{ij} \equiv V_1^{ij} - V_0^{ij} = \begin{cases} b_0 - b_{10}^{ij} V^{ij} + \Delta e_t^{ij} \mid IT = 1 \\ \Delta e_t^{ij} \mid IT = 0 \end{cases}. \tag{11}$$

[9] 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.

[10] 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.

To test equation (11) I estimate the following regression:

$$\Delta V^{ij} = b_0 + b_{0T} D_T^{ij} - V_0^{ij} \left(b_1 + b_{1T} D_T^{ij} \right) + u^{ij}, \tag{12}$$

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

B. Matching treatment and control groups

As is well known from the treatment effect literature, evaluating the effect of a treatment (in this case IT adoption) on the treated is made difficult by the fact that one does not observe the counterfactual (no treatment) for the treated group. One can infer this counterfactual from the behavior of an untreated group, but only if the treatment is randomly assigned or if one can simulate random assignment by selecting a control group whose probability of receiving the treatment, conditional on a set of observable variables, matches that of the treated (see Heckman and others, 1998, for a detailed discussion). I use propensity score matching, a form of matching on observables, to select the control group (Rosenbaum and Rubin, 1983; Smith and Todd, 2005). 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 $\{\Delta V^{ij}_{-1,g}; \Delta V^{ij}_{-1}\}$; and similar variables for the forecast level of these variables $\{f^{ij}_{0,g}; f^{ij}_{0}; \Delta f^{ij}_{-1,g}; \Delta f^{ij}_{-1}\}$. Table 1 reports the probit regression results. Targeting generally appears to be more strongly related to the forecaster variables relating to inflation $(V^{ij}_0, \Delta V^{ij}_{-1}, f^{ij}_0, \Delta f^{ij}_{-1})$, which suggests that controlling for endogeneity is likely to be particularly important for a study, such as this one, that focuses on the behavior of inflation forecasts. However, the overall fit of the regression is relatively low (.07), in part reflecting the large and diverse sample. The groups of control observations are 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.[11] 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.[12]

[Table 1 about here]

[Figure 1 about here]

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 the 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.

[11] 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).

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

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.

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 the sample—both are selected). Controls are then selected from this limited set of observations according to the propensity score (matching with replacement).

The appendix provides a full account of the matching algorithms used in the paper. Table 2 provides summary statistics (means with standard deviations in parentheses) for the treatment and three control groups. The treatment and three control groups are detailed in Table 3.

[Table 2 about here]

[Table 3 about here]

C. Dealing with mean reversion

From equations (11) and (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 $V_0^{ij} = \frac{1}{\left(\alpha^j(0)+\beta^i\right)} + e_0^{ij}$, 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. More critically, $cov\left(V_0^{ij}, u^{ij}\right) < 0$, 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 the 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 (12).

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 $\left(\frac{1}{\left(\alpha^j(0)+\beta^i\right)}\right)$ 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.[13] 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 this should fulfill the relevancy and exogeneity requirements of a "good" instrument.[14]

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_0^{ij}. Using expected rather than actual inflation eliminates the impact of unexpected inflationary shocks that

[13] 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.

[14] 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.

are likely to be correlated with the transient component of forecast errors.[15]

These two variables $\{V^{ij}_{0,g}, f^{ij}_0\}$ constitute suitable instruments for V^{ij}_0. Then, since the "treatment" dummy D_T is assumed exogenous (conditional on the matching of a control group of observations with the treatment observations), suitable instruments for $D_T V^{ij}_0$ are $\{D_T V^{ij}_{0,g}, D_T f^{ij}_0\}$.

V. Results

A. Graphical results

As a first pass, Figure 2 plots the change in the mean absolute inflation forecast error (ΔV^{ij}) against the average prior error (V^{ij}_0) for observations from IT adoption episodes and from the full pool of control episodes.[16] The relationship appears to be negative for all episodes (due to mean reversion), but there is evidence for an additional negative effect (conditional on the initial forecast error) due to IT adoption, as predicted. The next section provides a formal econometric analysis of this effect, paying attention to endogeneity issues.

[Figure 2 about here]

B. Matching results

The results of estimating equation (12) are presented in Table 4. Columns I-III present results for the three matching methods (1-3 respectively). For each group, the first set of results suppresses the interaction effect, measuring only the levels (or unconditional) effect of IT adoption on forecast errors; the second estimates the full equation using OLS; the

[15] Since shocks to forecast accuracy should, at least in theory, lead forecasters to over- and underestimate inflation with equal probability, f^{ij}_0 should not be correlated with e^{ij}_0.

[16] A common support for V_0 is imposed for IT and non-IT adoption episodes. In addition, outliers (defined as those with an absolute change in the inflation forecast error $|\Delta V|$ in excess of 10 percent) are also dropped. The sample includes 197 observations from IT adoption episodes and 2,048 observations from control episodes. The results are similar if the sample is not truncated to exclude outliers.

third 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.[17]

[Table 4 about here]

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 the residuals (allowing arbitrary patterns of correlations within groups) 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. All results therefore report clustered standard errors.[18]

Table 4 provides extremely strong evidence for the conditional effect predicted by the model, but little evidence for an unconditional effect. The coefficient b_{1T} estimated by OLS is between 0.6 (column III) and 0.75 (column I), which is quantitatively significant. The estimates are also highly statistically significant, at least at the 1% level (column I) and generally at the 0.1% level (columns II and III). The IV results 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

[17] 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.

[18] For Table 4, the number of clusters is 58 (panel A), 63 (panel B) and 23 (panel C). Note that non-clustered standard error estimates (either assuming iid errors or robust estimates) are lower, as one would expect.

highly significant as well as having a much higher point estimate.[19] The unconditional effect is not generally statistically significant although it carries a negative sign in all specifications.

C. Robustness checks

Table 5 provides some robustness checks, replicating Table 4 for two placebo datasets. Panel A replicates Table 4 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.[20] There is no statistically significant evidence for a conditional or unconditional "effect" of the placebo IT adoption variable. Panel B undertakes the same exercise shifting the adoption window *back 12 months*. Again, there is no evidence of an "effect" from the placebo.

[Table 5 about here]

Note that since the sample size is greater than that for the genuine data in this latter 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 5 therefore provides strong evidence that the results in Table 4 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.[21]

[19] 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). Test statistics and associated p-values or critical values are shown at the bottom of Table 4.

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

[21] The results from Panel B additionally suggest that IT adoption involves a one-time transparency gain rather than a gradual process that delivers additional gains in the period following IT adoption. This also provides some backing for the idea that the effect identified in the data derives from increased transparency rather than credibility, since the latter is likely to accrue over time whereas the former could be instantaneous.

Table 6 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 more pronounced 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 4, Table 6 presents results for three trimmed samples in which some controls with high values for V_0^{ij} have been dropped.

[Table 6 about here]

In Panel A the 5 percent of observations with the highest values for V_0^{ij} (all controls) are dropped from the dataset. In Panel B 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 column III, where the estimated effect is no longer statistically significant. On the other hand, the IV estimates of b_{1T} are strengthened in terms of both point estimates and statistical significance.[22] Finally, Panel C presents 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 column III (the "best country"

[22] 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.

control group).[23]

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).

VI. 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

[23] Even in this case the p–value associated with the IV estimates (.122) indicates borderline statistical significance. The reason for the large effect in this case is that Venezuela accounts for more than 20 percent of the total (weighted) controls under the "best country" matching algorithm.

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' forecasting advantage over real variables is likely far lower than for inflation.[24] A second area of further research is in establishing what dimensions of transparency associated with IT could account for the apparent effect on forecasters' accuracy. Crowe and Meade (2008) provide some evidence that aspects of transparency associated with the central bank's openness concerning its own data and forecasts (economic and operational transparency) tend to be most strongly associated with the private sector making greater use of public information rather than private signals. However, further work is needed to establish whether these results hold in the specific case

[24] 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.

of IT adoption.

Appendix: 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 IV $\left\{V_{0,g}^{ij}, V_0^{ij}, \Delta V_{-1,g}^{ij}, \Delta V_{-1}^{ij}, f_{0,g}^{ij}, f_0^{ij}, \Delta f_{-1,g}^{ij}, \Delta f_{-1}^{ij}\right\}$ and taking the fitted probability. The total number of observations used to estimate *ps* is 2,141.[25] 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)

[25] The probit regression has a χ^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.

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.*

References

Ang, Andrew, Geert. Bekaert and Min Wei (2005), Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better? NBER Working Paper 11538. Cambridge, MA: National Bureau of Economic Research.

Ball, Laurence and Niamh Sheridan (2004), "Does inflation targeting matter?" in Bernanke and Woodford (eds.), *The Inflation Targeting Debate*. Chicago: University of Chicago Press.

Berg, Claes (2005), "Experience of inflation-targeting in 20 countries," *Sveriges Riksbank Economic Review* 2005 (1), pp. 20-47.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), "How much should we trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119, pp. 249-275.

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? Mimeo., 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.

Chortareas, Georgios, David Stasavage and Gabriel Sterne (2003), "Does Monetary Policy Transparency Reduce Disinflation Costs?" *The Manchester School* 71, pp. 521-540.

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

Crowe, Christopher and Ellen E. Meade (2007), "Evolution of Central Bank Governance Around the World," *Journal of Economic Perspectives* 21, pp. 69-90.

Crowe, Christopher and Ellen E. Meade (2008), "Central Bank Independence and Transparency: Evolution and Effectiveness," *European Journal of Political Economy*, forthcoming.

Eijffinger, Sylvester and Petra Geraats (2006), "How Transparent are Central Banks?" *European Journal of Political Economy* 22(1), pp. 1-21.

Faust, Jon and Dale Henderson (2004), "Is Inflation Targeting Best-Practice Monetary Policy?" *Federal Reserve Bank of St. Louis Review* 86 (4) pp. 117-144.

Fry, Maxwell, DeAnne Julius, Lavan Mahadeva, Sandra Roger and Gabriel Sterne (2000), *Key issues in the choice of monetary policy framework*. In Mahadeva and Sterne (eds., 2000).

Geraats, Petra (2002), "Central Bank Transparency," *Economic Journal* 112, pp. F532-565.

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.

Heckman, J., H. Ichimura and E. Todd, 1998. "Matching as an econometric evaluation estimator." *Review of Economic Studies* 65, pp. 261-294.

Johnson, David (2002), "The effect of inflation targeting on the behavior of expected inflation: evidence from an 11 country panel," *Journal of Monetary Economics* 49, pp. 1521-1538.

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 29 October. Mimeo: Bank of England.

Kuttner, Kenneth and A. Posen (1999), "Does talk matter after all? Inflation targeting and central bank behavior," Staff Report 88, Federal Reserve Bank of New York.

Leuven, Edwin and Barbara Sianesi (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." http://ideas.repec.org/c/boc/bocode/s432001.html. Version 3.0.0.

Lin, Shu and Haichun Ye (2007), "Does inflation targeting really make a difference? Evaluating the treatment effect of inflation targeting in seven industrial countries." *Journal of Monetary Economics* 54, pp. 2521-2533.

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

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 8397.

Morris, Stephen. and Hyun Song Shin (2002), "Social Value of Public Information," *American Economic Review* 92 (5), pp. 1521-1534.

Ottavianni, M. and P. N. Sørensen (2006), "The Strategy of Professional Forecasting," *Journal of Financial Economics* 81, pp. 441-466.

Petursson, Thórarinn. (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 WP/05/163.

Romer, Christine and David Romer (2000), "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review* 90 (3), pp. 429-457.

Rosenbaum, R. and B. Rubin (1983). "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70, pp. 41-55.

Smith, Jeffrey A. and Petra E. Todd (2005), "Does matching address Lalonde's critique of nonexperimental estimators?", *Journal of Econometrics* 125 (2), pp. 305-353.

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* 20 (4), pp. 518-29.

Stock, James H. and Motohiro Yogo (2002), Testing for Weak Instruments in Linear IV Regression. Technical Working Paper 284, National Bureau of Economic Research.

Svensson, Lars (1999), "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics* 43, pp. 607-654.

Svensson, Lars (2006), "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, not Con," *American Economic Review* 96 (1), pp. 448-452.

Vega, Marco and Diego Winkelried (2005), "Inflation Targeting and Inflation Behavior: A Successful Story?" *International Journal of Central Banking* 1 (3), pp. 153-75.

Tables and Figures

Table 1. Probit estimate of propensity score

Dependent Variable	IT
V_0^{ij}	$-.159^{***}$
	(.0356)
$V_{0,g}^{ij}$	$-.0540^{*}$
	(.0303)
f_0^{ij}	$.0704^{***}$
	(.0170)
$f_{0,g}^{ij}$	$.0560^{*}$
	(.0320)
ΔV_{-1}^{ij}	$.140^{***}$
	(.0313)
$\Delta V_{-1,g}^{ij}$	$-.113^{***}$
	(.0254)
Δf_{-1}^{ij}	$-.116^{***}$
	(.0286)
$\Delta f_{-1,g}^{ij}$	.0417
	(.0353)
Observations	$2,141$
Pseudo-R^2	.0744
LR χ^2 (8)	87.60^{***}

Constant term included but not reported. Standard errors reported in parentheses.
Significance levels denoted as: ***: 1 percent; **: 5 percent; *: 10 percent.

Table 2. 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

Table 3. [attached after Table 6]

Table 4. Regression Results

ΔV^{ij}	I			II			III		
Control Group	Nearest Neighbor (with replacement)			Nearest Neighbor (without replacement)			"Best Country" (with replacement)		
Model	Levels	OLS	2SLS	Levels	OLS	2SLS	Levels	OLS	2SLS
IT	−1.16*	.200	.831**	−.961	−.0653	.589	−.827	−.389	.472
	(.666)	(.423)	(.421)	(.729)	(.489)	(.462)	(1.26)	(.947)	(.735)
V_0^{ij}		−.153	.0382		−.286**	−.0978		−.365***	−.163
		(.124)	(.231)		(.112)	(.211)		(.107)	(.223)
$IT \times V_0^{ij}$		−.752†††	−.944†††		−.637†††	−.820†††		−.598***	−.800***
		(.154)	(.238)		(.152)	(.231)		(.190)	(.284)
Adv.	.0742	−.803	−.306	.309	−.973	−.448	.453	−1.36	−.634
	(.690)	(.528)	(.419)	(.753)	(.619)	(.485)	(1.30)	(1.25)	(1.04)
Const.	.193	1.16*	.308	−.112	1.54**	.641	−.311	2.13	.938**
	(.864)	(.648)	(.611)	(.979)	(.765)	(.675)	(1.80)	(1.66)	(1.41)
$F-stat$	2.47*	44.4†††	22.8†††	2.37	39.7†††	20.0†††	2.22	26.8†††	16.6†††
R^2	.0545	.216	.178	.0366	.225	.190	.0265	.246	.203
id test (p-value)			172.9			172.7			171.1
			(.000)			(.000)			(.000)
over-id test (p-value)			1.42			1.99			1.43
			(.493)			(.370)			(.488)
CD statistic (critical value)			55.5			55.5			54.8
			(16.9)			(16.9)			(16.9)

Significance levels denoted as: ††† 0.1 percent; ***: 1 percent; **: 5 percent; *: 10 percent.
id test: Anderson canon. corr. LR statistic; over-id test: Hansen J statistic;
CD: Cragg-Donald statistic (note that this is based on homoskedastic SEs; critical value is based on 2SLS size, 5% sig. level, max. size 10%).

Table 5. Placebo Regressions

ΔV^{ij}	I			II			III		
Control Group	Nearest Neighbor (with replacement)			Nearest Neighbor (without replacement)			"Best Country" (with replacement)		
Model	Levels	OLS	2SLS	Levels	OLS	2SLS	Levels	OLS	2SLS
Panel A: 12 Months Before									
IT	−.648	.617	.0828	−1.05	−.180	−.0506	−.171	−.681	.434
	(.495)	(.556)	(.868)	(.652)	(.646)	(.830)	(.412)	(.492)	(1.18)
V_0^{ij}		.152	.360		−.0441	.254*		−.737†††	.554
		(.235)	(.246)		(.140)	(.152)		(.111)	(.810)
$IT \times V_0^{ij}$		−.457	−.209		−.352	−.251		.356	−.352
		(.284)	(.313)		(.239)	(.260)		(.237)	(.626)
$Adv.$	−.322	−.227	.471	−.638	−1.01*	−.00165	.111	−.874**	.798
	(.423)	(.419)	(.557)	(.544)	(.592)	(.596)	(.329)	(.399)	(1.10)
$Const.$	.308	−.163	−1.08	.812	1.13	−.400	−.306	1.54***	−1.67
	(.505)	(.701)	(.768)	(.708)	(.834)	(.762)	(.342)	(.497)	(1.97)
$F-stat$	2.33*	1.44	.64	1.35	1.37	.99	.12	17.5†††	.13
R^2	.0206	.0752	< 0	.0329	.0514	< 0	.0063	.225	< 0
clusters		57			62			20	
obs. (weighted)		272			272			272	
Panel B: 12 Months After									
IT	−.115	.218	.705	−.340	−.0737	.808	.524	.178	.300
	(.395)	(.455)	(.736)	(.435)	(.479)	(.765)	(.415)	(.532)	(.760)
V_0^{ij}		.0312	.375		−.0129	.494		−.395	−.326
		(.187)	.241		(.222)	(.349)		(.284)	(.312)
$IT \times V_0^{ij}$		−.273	−.588		−.229	−.820		.153	.0742
		(.494)	(.674)		(.503)	(.706)		(.577)	(.699)
$Adv.$	−.219	−.194	.0715	−.470	−.481	−.0656	.294	−.00572	.0468
	(.403)	(.442)	(.490)	(.450)	(.457)	(.514)	(.391)	(.462)	(.515)
$Const.$	.296	.238	−.390	.623	.647	−.305	−.554*	.202	.0698
	(.380)	(.475)	(.603)	(.472)	(.482)	(.706)	(.296)	(.576)	(.681)
$F-stat$	.20	.42	.86	.76	.87	.87	1.41	2.04	1.23
R^2	.0057	.0114	< 0	.0224	.0264	< 0	.0649	.131	.130
clusters		69			73			26	
obs. (weighted)		352			352			352	

Significance levels denoted as: ††† 0.1 percent; ***: 1 percent; **: 5 percent; *: 10 percent.

Table 6. Robustness Checks

ΔV^{ij}	I			II			III		
Control Group	Nearest Neighbor (with replacement)			Nearest Neighbor (without replacement)			"Best Country" (with replacement)		
Model	Levels	OLS	2SLS	Levels	OLS	2SLS	Levels	OLS	2SLS
Panel A: $F\left(V_0^{ij}\right) \leq .95$									
IT	-1.53^{**}	$1.35^{\dagger\dagger\dagger}$	$3.04^{\dagger\dagger\dagger}$	-1.30^{*}	$.644^{*}$	2.35^{**}	-1.53	.561	3.02^{**}
	(.696)	(.371)	(.571)	(.711)	(.363)	(.924)	(1.42)	(.582)	(1.20)
V_0^{ij}		$.753^{*}$	$1.94^{\dagger\dagger\dagger}$		.0858	.948		.148	1.41^{**}
		(.416)	(.272)		(.352)	(.577)		(.760)	(.589)
$IT \times V_0^{ij}$		$-1.59^{\dagger\dagger\dagger}$	$-2.66^{\dagger\dagger\dagger}$		$-.950^{**}$	$-1.68^{\dagger\dagger\dagger}$		-1.03	$-2.07^{\dagger\dagger\dagger}$
		(.424)	(.265)		(.363)	(.522)		(.782)	(.618)
$Adv.$	$-.315$	$-.186$	.896	$-.0398$	$-.407$	.952	$-.274$	$-.579$	1.51
	(.751)	(.311)	(.571)	(.750)	(.337)	(.767)	(1.48)	(.457)	(1.56)
$Const.$	.744	$-.407$	$-2.84^{\dagger\dagger\dagger}$	.382	.451	-2.15	.724	.651	$-.323$
	(.910)	(.439)	(.843)	(.943)	(.436)	(1.38)	(2.03)	(.607)	(2.20)
$F-stat$	2.98^{*}	$84.1^{\dagger\dagger\dagger}$	$42.8^{\dagger\dagger\dagger}$	2.62^{*}	$66.8^{\dagger\dagger\dagger}$	$30.2^{\dagger\dagger\dagger}$	1.92	$53.0^{\dagger\dagger\dagger}$	$20.1^{\dagger\dagger\dagger}$
R^2	.106	.415	.145	.0658	.217	< 0	.0710	.210	< 0
obs (clusters)		307 (54)			316 (62)			306 (21)	
Panel B: $V_0^{ij} \leq \max\left(V_0^{ij} \mid IT = 1\right)$									
IT	-1.26^{**}	$.992^{***}$	$3.30^{\dagger\dagger\dagger}$	-1.31^{**}	$.836^{*}$	$3.04^{\dagger\dagger\dagger}$	-1.41	.570	2.92^{***}
	(.518)	(.456)	(.883)	(.630)	(.464)	(.828)	(1.16)	(.626)	(.998)
V_0^{ij}		.395	$2.17^{\dagger\dagger\dagger}$		.236	1.57^{***}		.137	1.56^{**}
		(.467)	(.677)		(.444)	(.558)		(.820)	(.612)
$IT \times V_0^{ij}$		-1.25^{***}	$-2.88^{\dagger\dagger\dagger}$		-1.09^{**}	$-2.27^{\dagger\dagger\dagger}$		-1.02	$-2.27^{\dagger\dagger\dagger}$
		(.466)	(.638)		(.444)	(.526)		(.839)	(.641)
$Adv.$	$-.0332$	$-.294$	1.01	$-.0551$	$-.314$	1.22^{*}	$-.143$	$-.555$	1.15
	(.585)	(.327)	(.663)	(.681)	(.335)	(.652)	(1.22)	(.434)	(1.25)
$Const.$	.345	.0275	-3.17^{***}	.404	.197	-3.03^{***}	.537	.626	-2.85
	(.651)	(.558)	(1.23)	(.823)	(.565)	(1.14)	(1.65)	(.639)	(1.77)
$F-stat$	3.32^{**}	$75.1^{\dagger\dagger\dagger}$	$28.7^{\dagger\dagger\dagger}$	2.89^{*}	$73.6^{\dagger\dagger\dagger}$	$25.2^{\dagger\dagger\dagger}$	2.07	$54.6^{\dagger\dagger\dagger}$	$24.8^{\dagger\dagger\dagger}$
R^2	.0920	.350	< 0	.0804	.279	< 0	.0720	.236	< 0
obs (clusters)		304 (54)			308 (61)			298 (21)	
Panel C: Venezuela dropped									
IT	$-.529$	.241	$.746^{**}$	$-.390$	.0618	$.684^{**}$	$-.106$	.106	.470
	(.373)	(.224)	(.322)	(.376)	(.217)	(.337)	(.773)	(.312)	(.248)
V_0^{ij}		$-.560^{\dagger\dagger\dagger}$	$-.285$		$-.707^{\dagger\dagger\dagger}$	$-.377$		$-.757^{\dagger\dagger\dagger}$	$-.600^{\dagger\dagger\dagger}$
		(.0757)	(.250)		(.0811)	(.264)		(.0891)	(.223)
$IT \times V_0^{ij}$		$-.290^{\dagger\dagger\dagger}$	$-.594^{**}$		$-.144^{*}$	$-.500^{*}$		$-.0859$	$-.279$
		(.0811)	(.246)		(.0849)	(.257)		(.0616)	(.180)
$Adv.$	.729	$-.274$	$-.0579$	$.898^{*}$	$-.283$	.00261	1.19	$-.211$	$-.0592$
	(.452)	(.243)	(.255)	(.459)	(.240)	(.251)	(.846)	(.362)	(.295)
$Const.$	$-.735$	$.765^{**}$	.226	$-.949^{*}$	$.950^{***}$	.255	-1.37	.858	.502
	(.471)	(.304)	(.424)	(.486)	(.298)	(.443)	(1.11)	(.499)	(.343)
$F-stat$	2.41^{*}	$91.7^{\dagger\dagger\dagger}$	$29.2^{\dagger\dagger\dagger}$	2.59^{*}	$98.5^{\dagger\dagger\dagger}$	$31.0^{\dagger\dagger\dagger}$	2.58	$173^{\dagger\dagger\dagger}$	$30.9^{\dagger\dagger\dagger}$
R^2	.138	.771	.723	.138	.766	.709	.152	.803	.786
obs (clusters)		305 (53)			302 (57)			294 (20)	

Table 3. 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 3 (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

Figure 1. Distribution of Propensity Score in Treatment and Control Groups

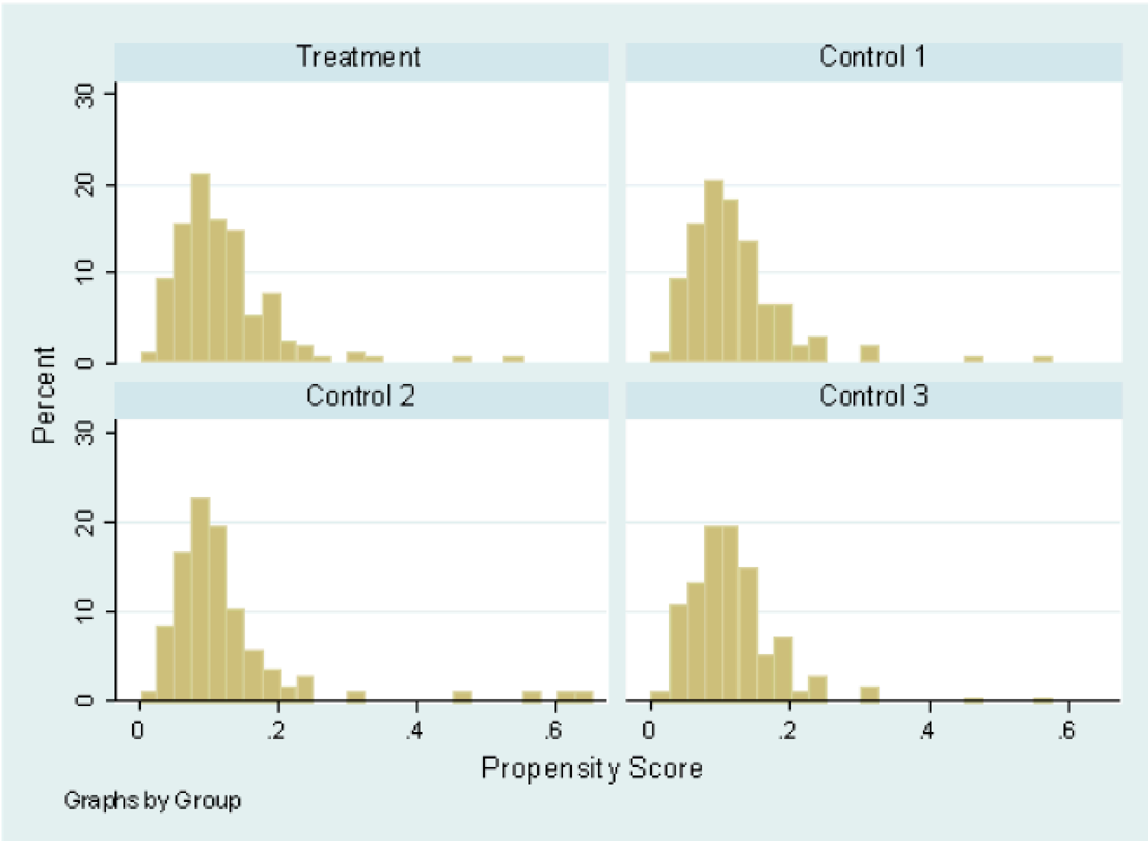

Figure 2. Change in Inflation Forecast Error