Inflation Targeting and Inflation Behavior: A Successful Story?

Marco Vega and Diego Winkelried

17. June 2004

Inflation Targeting and Inflation Behavior: A Successful Story?*

Marco Vega[a] and Diego Winkelried[b]
[a]London School of Economics and Central Bank of Peru
[b]St. Edmund's College, University of Cambridge

This paper estimates the effects of inflation targeting (IT) adoption over inflation dynamics using a wide control group. We contribute to the current IT evaluation literature by considering the adoption of IT by a country as a treatment, just as in the program evaluation literature. Hence, we perform propensity score matching to find suitable counterfactuals to the actual inflation targeters. We find out that IT has helped in reducing the level and volatility of inflation in the countries that adopted it. This result is robust to alternative definitions of treatment and control groups.

JEL Codes: C50, E42, E52.

Since the early 1990s, an increasing number of central banks have adopted an inflation targeting (IT) regime as a monetary policy framework, spurring research on the benefits of such a policy scheme.[1] Theoretical work suggests that the sound implementation of an IT regime delivers "optimal" equilibrium, in the sense of anchoring inflation around a target with relatively low inflation and,

*We would like to thank the editor for his support and an anonymous referee for suggestions that have improved the paper. We are grateful to Charles Goodhart and seminar participants at the Central Bank of Peru and the University of Cambridge for their valuable comments. Any remaining errors and the opinions herein are our own. Corresponding author: Winkelreid: St. Edmund's College, University of Cambridge, CB3 0BN, Cambridge, United Kingdom; e-mail: dw295@cam.ac.uk; phone (++ 44) 7765 574 055. Other author contact: Vega: Central Bank of Peru, Jr. Miroquesada 441, Lima 1, Peru; e-mail: m.a.vega@lse.ac.uk.

[1]To date, twenty-one countries follow an explicit IT framework and other countries are considering its adoption. See, for example, Truman (2003) and Pétursson (2004).

if "flexible," low output volatility.[2] However, whether IT has led to a superior monetary policy performance or has induced macroeconomic benefits in the countries that adopted it (ITers henceforth) is definitely an empirical matter.

A fundamental element to be taken into account in any empirical work is the appropriate success measure. From the onset, we make a distinction between relative and absolute success criteria. Absolute criteria refer to a type of policy evaluation that accounts gains in macroeconomic outcomes without reference to alternative policies that might have achieved the same outcomes. Relative criteria, on the other side, evaluate a particular policy framework in comparison to others to assess whether the former is superior.

The verdict of the absolute criteria about IT success is overwhelmingly one:[3] "IT has been beneficial." To our knowledge, there is no empirical work that has found that IT has delivered worse outcomes in comparison to preadoption ones.[4] On the other hand, the relative criteria have not yet yielded a clear-cut conclusion. This approach would attempt to answer questions like *does inflation targeting make a difference?* or *does inflation targeting matter?* From a policy evaluation view, this is the relevant criteria at which we need to look.

Our goal is to evaluate the behavior of inflation dynamics brought about by the adoption of IT. We do so by studying three measures that distinguish inflation dynamics: mean, variance, and persistence. Key questions emerge from the study of these measures.

First, IT has been adopted by countries either to credibly disinflate (or converge) or, as asserted by some authors, to lock in the gains obtained from episodes of disinflation. Would countries have done better or worse had they adopted any other regime?

Second, it is generally stated that inflation uncertainty results from factors exogenous to the scope of the transmission mechanism of monetary policy (terms of trade or supply shocks, for instance) as well as from monetary policy shocks. In this sense, inflation can be made less uncertain up to the limits set out by the amount of exogenous uncertainty. Modern monetary policy practice, whether IT or

[2] See, for example, Svensson (2000).

[3] See, for example, Bernanke et al. (1999), Mishkin and Schmidt-Hebbel (2002), and Corbo, Landarretche, and Schmidt-Hebbel (2002).

[4] See Neumann and von Hagen (2002) for a recent empirical survey.

not, hinges precisely on making monetary policy more predictable and hence less uncertain. Once again, a fair question for a country that adopted IT is whether inflation uncertainty has fallen more or less in comparison to the counterfactual situation of not having adopted IT.

Last, the theory of IT emphasizes that the overall features of the framework are built upon the pillar of credibility. Credibility is understood as the ability the central bank has to anchor medium- to long-run expectations, to avoid expectation traps that may render persistently high or low inflation rates. On the other hand, "flexible" IT implies that shocks that drive inflation away from the target should revert at a pace that does not harm real activity. Hence, the speed of adjustment seems to depend on the degree of flexibility.[5] Too fast an adjustment is equivalent to a strict IT, likely in situations whereby the central bank needs to gain or strengthen credibility. When the adjustment is slow, a more flexible IT is in place. In the fast-adjustment case, undue real volatility might emerge, whereas in the slow-adjustment case either credibility is strong enough that the central bank can reap some benefits of flexibility or the nominal anchor is lost and the inflation falls to the expectation trap.

Thus, the effects of IT adoption on persistence are ambiguous. More persistence can result from successful flexible ITers or unsuccessful ITers not gaining credibility. Once more, what does an empirical evaluation of IT over persistence tell about the adopting ITers?

In recent years, a growing body of literature has provided insights on the empirical assessment of IT. Corbo, Landarretche, and Schmidt-Hebbel (2002), for instance, compare policies and outcomes in fully fledged IT countries to two groups, potential ITers and non-ITers. They find that sacrifice ratios were lower in ITers, that IT countries have reduced inflation forecast errors, and that inflation persistence has declined strongly among ITers.

Johnson (2002), by comparing five ITers to six non-ITers, all of them in industrialized economies, finds that the period after the announcement of IT is associated with a statistically significant reduction in the level of expected inflation. Also, he finds that IT has not reduced absolute average forecast errors in targeting countries relative to those in nontargeting countries. However, ITers did avoid

[5]See Svensson (1999).

even larger forecast errors than those that would have occurred in the absence of IT.

On the other hand, Neumann and von Hagen (2002) consider a group of six industrial IT countries and three non-IT countries and perform an event study to quantify the response of inflation and long-run as well as short-run interest rates to supply shocks (increases in the world oil price in 1978–79 and in 1998–99[6]). They find that the effect of IT is not significantly different from zero for average inflation, but it is for interest rates, meaning a gain in credibility among ITers.

Pétursson (2004) analyzes a bigger sample (twenty-one ITers) that includes developing economies. He evaluates the performance of a set of macroeconomic outcomes using a dummy variable for pre- and post-IT periods on a country basis and finds that IT has been beneficial to reduce the level, persistence, and variability of inflation.[7] However, the technique offered by this study does not tackle the fundamental question of relative performance. Its contribution hinges in giving a clear and robust account for the evidence of the absolute benefits of IT and corroborates previous findings on this line.

Levin, Natalucci, and Piger (2004) study inflation persistence using five industrial ITers that are compared to seven industrial non-ITers. The study performs univariate regressions on inflation for each country and finds that inflation persistence is estimated to be quite low within ITers, whereas the unit root hypothesis cannot be rejected for non-ITers. Levin and Piger (2004), on the other hand, in a similar empirical framework with twelve industrial countries, allow for structural breaks and find that inflation, in general, exhibits low persistence.[8] They also suggest that IT does not seem to have had a large impact on long-term expected inflation for a group of eleven emerging market economies.

[6]This type of shock creates a dilemma because it implies more inflation coupled with a downturn of economic activity.

[7]There are other studies that provide mixed evidence about inflation persistence. Benati (2004) and Levin, Natalucci, and Piger (2004) find that inflation has become less persistent within the OECD and especially IT countries.

[8]These results confirm those of Benati (2004), which studies inflation dynamics in twenty OECD countries.

Finally, Ball and Sheridan (2005) provide evidence on the irrelevance of IT. They look at seven OECD countries that adopted IT in the early 1990s and thirteen countries that did not. They claim that ITers that reduced higher-than-average inflation rates toward equilibrium levels were merely reflecting *regression to the mean* and not a proper effect of IT. Once they control for regression to the mean, they conclude that IT did not improve macroeconomic performance. In their words, "Just as short people on average have children who are taller than they are, countries with unusually high and unstable inflation tend to see these problems diminish, regardless of whether they adopt inflation targeting."

In our view, rather than challenging the previous evidence and beliefs about IT effects, the crucial point of the claim made in Ball and Sheridan (2005) is methodological. If there is an ITer with poor performance before IT, then it should be compared with a non-ITer with equally poor initial performance. Otherwise, the targeting effect would be overstated. Here, we hinge precisely on this matter of comparability.

Following Johnson (2002) and Ball and Sheridan (2005), we use a difference-in-difference estimator approach to evaluate the effects on key measures of inflation dynamics resulting from IT adoption. As we argue later, the previous studies on this issue may suffer from sample selection bias (a few industrialized countries, for instance) and, importantly, select counterfactuals for the ITers in an arbitrary fashion. Our contribution is twofold: First, we use all the twenty-three IT experiences so far, the *widest possible control group* of non-ITers (eighty-six countries), and different possible dates of IT adoption. With this, we understand IT as an alternative monetary policy framework worldwide, for both industrialized and developing economies. Second, we interpret the IT adoption as a "natural experiment," so we seek to reestablish the conditions of a randomized experiment where the IT adoption mimics a *treatment*. This naturally leads us to perform *propensity score matching* as an alternative to the widely used regression approach. In a nutshell, we seek to overcome the aforementioned methodological limitations by letting the data select the controls for ITers.

The rest of the paper is organized as follows. In section 1 we briefly describe the propensity score and matching techniques for evaluation; in section 2 we discuss some empirical issues regarding the robustness of our results and present the inflation outcomes to be evaluated; in section 3 we show our main findings, while section 4 concludes and provides some avenues for future research.

1. Methodology

As mentioned, we use microeconometric techniques usually applied in quasi-experimental contexts, borrowed from the program evaluation literature. To be consistent with this literature in this section we may refer to the adoption of IT as *treatment*, to the ITers as the *treated* group, and to the non-ITers as the *control* group.

1.1 The Fundamental Problem

Let D be a binary indicator that equals one if a country has adopted IT and zero otherwise. Also, let Y_t^1 denote the value of a certain outcome in period t if the country has adopted the IT regime and Y_t^0 if not. Given a set of observable country attributes $\boldsymbol{X}$, the average effect of being an ITer on Y_t is[9]

$$\xi = E\big[\big(Y_t^1 - Y_t^0\big)\,\big|\,\boldsymbol{X}, D = 1\big] = E\big[Y_t^1\,\big|\,\boldsymbol{X}, D = 1\big] - E\big[Y_t^0\,\big|\,\boldsymbol{X}, D = 1\big]. \tag{1}$$

It is clear from (1) that we face an identification problem since $E[Y_t^0 \mid \boldsymbol{X}, D = 1]$ is not observable. It is convenient to rewrite (1) in a slightly different way, closer to what we actually use in our empirical work. Suppose that IT was adopted in period k. Then, for $t > k > t'$, (1) is equivalent to

$$\xi = E\big[\big(Y_t^1 - Y_{t'}^0\big)\,\big|\,\boldsymbol{X}, D = 1\big] - E\big[\big(Y_t^0 - Y_{t'}^0\big)\,\big|\,\boldsymbol{X}, D = 1\big]. \tag{2}$$

This way of representing ξ allows us to exploit the panel data nature of the sample, and hence to control for fixed factors that could be correlated with the outcomes (for instance, most developed countries having less volatile inflation rates).

A common approach to estimate the expectation $E\big[\big(Y_t^0 - Y_{t'}^0\big)\big|\boldsymbol{X}, D = 1\big]$ is to replace it with the observable average outcome

[9]The quantity ξ refers to what is defined in the literature as the *average treatment effect on the treated*, i.e., the average effect of IT only across those countries that adopted the regime.

in the untreated state $E[(Y_t^0 - Y_{t'}^0) \mid \boldsymbol{X}, D = 0]$.[10] However, this could result in biased estimates of ξ from two sources.[11] The first arises from the presence of ITers in the sample that are *not comparable* with non-ITers and vice versa. The second is due to *different distributions* of $\boldsymbol{X}$ between the treated and the control groups, which is usual in nonrandomized samples (like a data set of countries). Fortunately, matching methods deal with these shortcomings.

1.2 Matching Methods

Matching techniques seek to eliminate the aforementioned biases by pairing ITers with non-ITers that have similar observed characteristics. The goal is to estimate a suitable *counterfactual* for each ITer, and hence attribute the difference between the ITer's outcome and that of a matched counterfactual to the treatment.

The key identifying assumption is that we can reestablish the conditions of a randomized experiment (that is, random assignment of $\boldsymbol{X}$) when no such data are available. This means that countries with the same observable characteristics face a randomized experiment as to whether they receive the treatment or not. A direct implication of such an assumption is that if there are any omitted variables that explain whether the treatment is received, and if these variables are important for outcomes, then identification may not be achieved.

1.2.1 The Propensity Score

Usually, determining along which dimension to match the countries or what type of weighting scheme to use is a difficult task. Rosenbaum and Rubin (1983) reduce the dimensionality of this problem by suggesting that the match can be performed on the basis of a single index that summarizes all the information from the observable covariates. This index, the *propensity score*, is the probability of treatment conditional on observable characteristics,

$$p(\boldsymbol{X}) = E\left[D \mid \boldsymbol{X}\right] = \Pr\left(D = 1 \mid \boldsymbol{X}\right), \tag{3}$$

and should satisfy the *balancing hypothesis*, which states that observations with the same propensity score must have the same

[10] See, for instance, Johnson (2002) and Ball and Sheridan (2005).

[11] See Heckman et al. (1998).

distribution of $\boldsymbol{X}$ independently of the treatment status.[12] Hence, equation (2) can be rewritten as

$$\xi = E\left[(Y_t^1 - Y_{t'}^0) \mid p(\boldsymbol{X}), D = 1\right] - E\left[(Y_t^0 - Y_{t'}^0) \mid p(\boldsymbol{X}), D = 1\right]. \tag{4}$$

The noncomparability bias can be eliminated by only considering countries within the *common support*, the intersection on the real line of the supports of the distributions $\{p(\boldsymbol{X}) \mid D = 1\}$ and $\{p(\boldsymbol{X}) \mid D = 0\}$. The bias from different distributions of $\boldsymbol{X}$ is eliminated by reweighing the non-ITer observations.

Estimating the propensity score is straightforward, as any probabilistic model suits (3). For instance, we can adopt the parametric form $\Pr(D_i = 1 \mid \boldsymbol{X}_i) = F(h(\boldsymbol{X}_i))$, where $F(.)$ is the logistic cumulative distribution (a logit). However, two points are to be handled with care. First, the estimation requires choosing a set of conditioning variables $\boldsymbol{X}$ that are not influenced by the adoption of the IT regime. Otherwise, the matching estimator will not correctly measure the treatment effect, because it will capture the (endogenous) changes in the distribution of $\boldsymbol{X}$ induced by the IT adoption. For this reason, the $\boldsymbol{X}$ variables should measure country attributes *before* the treatment.[13] Second, the model selection, i.e., the form of $h(\boldsymbol{X}_i)$, can be used to test the balancing hypothesis. Dehejia and Wahba (2002) suggest using a polynomial according to the following steps:

1. Start with a parsimonious logit specification (i.e., $h(\boldsymbol{X}_i)$ linear).

2. Stratify all observations on the common support such that estimated propensity scores within a stratum for treated and control countries are close. For example, start by dividing observations into strata of equal score range $(0-0.2, \ldots, 0.8-1)$.

3. For each interval, test whether the averages of $\boldsymbol{X}$ of treated and control units do not differ. If covariates are balanced between

[12] Rosenbaum and Rubin (1983) show that the conditions $D \perp \{Y^1, Y^0\} \mid \boldsymbol{X}$ and $0 < p(\boldsymbol{X}) < 1$ together (*strong ignorability of the treatment*) are sufficient to identify the treatment effect. In practice, we require a weaker and *testable* condition of ignorability for identification: *conditional mean independence*, $E[Y^0 \mid \boldsymbol{X}, D] = E[Y^0 \mid \boldsymbol{X}]$ and $E[Y^1 \mid \boldsymbol{X}, D] = E[Y^1 \mid \boldsymbol{X}]$.

[13] However, even these variables could be influenced by the program through the effects of expectations.

these groups for all strata, the specification satisfies the balancing hypothesis.[14] If the test fails in one interval, divide it into smaller strata and reevaluate.

4. If a covariate is not balanced for many strata, a less parsimonious specification of $h(\boldsymbol{X}_i)$ is needed. This can be achieved by adding interaction and/or higher-order terms of the covariate.

It is important to emphasize that the role of the propensity score is to reduce the dimensionality of the matching; it does not necessarily convey a behavioral interpretation.[15] Indeed, the logit regressions do not seek to find the determinants that made a central bank adopt an IT regime, but to characterize and summarize the economic state in which the ITers began to implement the regime. The difference is subtle but allows us to control for variables that, although useful to define the profile of a particular economy (importantly, relative to others), are not theoretically included in the central bank's decision to change the monetary policy regime.[16]

1.2.2 The Matched Estimator

Given the propensity score, there are various methods available for finding a counterfactual for ITer i.[17] Following Heckman, Ichimura, and Todd (1997 and 1998), we can compute a consistent estimator of the counterfactual by means of a kernel weighted average of outcomes. This approach not only has good statistical properties, but is also a convenient way to work with a sample of countries, as it could be difficult to find an actual non-ITer for each ITer. Let $\mathcal{C}$ denote the set of non-ITer countries whose propensity scores are over the region of the common support. The counterfactual of the outcome $Y^0_{i,t}$ is

$$\tilde{Y}^0_{i,t} = \frac{\sum_{j\in\mathcal{C}} K_b(p_j - p_i)Y^0_{j,t}}{\sum_{j\in\mathcal{C}} K_b(p_j - p_i)}, \tag{5}$$

[14] Actually, the specification satisfies the weaker version of conditional mean independence. See footnote 12.

[15] See Mishkin and Schmidt-Hebbel (2002) for an attempt to interpret a cross-sectional logit of the IT adoption in behavioral terms.

[16] In our empirical application we use a wide set of country attributes in modeling the propensity score to reduce the odds of an omitted variable problem and to minimize possible identification problems (see section 2.2).

[17] See Smith and Todd (2005) for a review and examples.

where $K_b(z) = K\left(\frac{z}{b}\right)$ is a kernel function (with bandwidth parameter b) that weights the outcome of country i inversely proportionally to the distance between its propensity score value (p_i) and the one of the non-ITer j (p_j).

Having found the matched pairs of ITers and non-ITers, the treatment effect estimator for country i in period $t > k$ can be written as

$$\hat{\xi}_{i,t} = \left(Y_{i,t}^{1} - \frac{1}{k-1} \sum_{\tau=1}^{k-1} Y_{i,\tau}^{0} \right) - \left(\tilde{Y}_{i,t}^{0} - \frac{1}{k-1} \sum_{\tau=1}^{k-1} \tilde{Y}_{i,\tau}^{0} \right), \qquad (6)$$

where the pretreatment outcome $Y_{t'}^{0}$ has been replaced by the time averages of $Y_{i,\tau}^{0}$ and $\tilde{Y}_{i,\tau}^{0}$ before the treatment.[18] The estimator (6) has no analytical variance, so standard errors are to be computed by bootstrapping (i.e., resampling the observations of the control group). Finally, the average of all possible $\hat{\xi}_{i,t}$ constitutes an unbiased estimator of (2),

$$\hat{\xi} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{T_i} \sum_{t=1}^{T_i} \hat{\xi}_{i,t} \right), \qquad (7)$$

where N is the number of ITers in the sample and T_i is the number of years ITer i has been conducting its monetary policy under an IT regime.

2. Empirical Issues

Before presenting the propensity score estimations and the "inflation outcomes" to be used in our evaluation, it is convenient to briefly discuss some issues regarding the dates the various central banks adopted their IT regime, i.e., the period when treatment occurred.

2.1 Adoption Dates

In a number of cases, the exact IT adoption timing is unclear: authors and central banks use different criteria. To address this ambiguity

[18] Heckman, Ichimura, and Todd (1998) and Smith and Todd (2005) suggest using a weighted average of the pretreatment observations instead of a sole observation to control for possible outliers or trend effects. In (6) we have used a simple average (equal weights).

and for the sake of robustness, we use two possible adoption dates for each country. First, we consider dates when countries started some form of IT (*soft* IT), typically by simply announcing numerical targets for inflation or by stating that they were switching to IT. On the other hand, we use dates of *fully fledged* IT adoption, namely, an explicit IT adoption as publicized by central banks and implying numerical targets for inflation together with the absence of nominal anchors other than the inflation target.[19]

Our approach contrasts with previous studies as it considers that many developing-country ITers used a soft version of IT as a strategy to reduce inflation from two-digit to international levels;[20] once inflation reached a stable low level, their central banks would reinforce the regime, by abandoning other nominal anchors and committing exclusively to target inflation. For example, Chile may appear as an early IT adopter (1991) in other studies, but it ran exchange rate regimes not compatible with fully fledged IT until 1999. For Peru, authors such as Corbo, Landarretche, and Schmidt-Hebbel (2002) use a soft IT adoption date (1994), when the central bank announced an inflation target consistent with a money growth operational target, while Levin, Natalucci, and Piger (2004) use its fully fledged date (2002).

The year of IT adoption for developed economies is less controversial. In New Zealand, for instance, the beginning of IT can be dated as far as 1988 when a numerical target for inflation was announced in the government budget statement. Or, following Mishkin and Schmidt-Hebbel (2002), it can be dated back to 1990 when the first Policy Targets Agreement between the minister of finance and the governor of the Reserve Bank of New Zealand was published, specifying numerical targets for inflation and the dates by which they were to be achieved. In 1991, a target range of 0 to 2 percent for 1993 was announced.[21]

In the case of Sweden, we follow Ball and Sheridan (2005) for our fully fledged classification given that the first announced inflation target was 2 percent for 1995 even though the Riksbank announced

[19]This information is available from the various central bank's web sites.

[20]See Fraga, Goldfajn, and Minella (2003) for a comprehensive survey of IT in developing countries.

[21]The upper bound of this range was changed to 3 percent in the 1996 Policy Target Agreement.

its shift to IT during 1993. For Canada, the first target range was announced in 1991. In 1993, a range of 1 to 3 percent was established for 1994 onward.

In table 1 we compare adoption dates among six different studies and provide our two possible adoption dates. The column labeled "Class. 1" refers to the soft IT adoption dates, while "Class. 2" accounts for fully fledged IT adoption. In six cases we have more than a three-year difference between both dates: Chile (eight years), Colombia (four years), Israel (five years), Mexico (four years), Peru (eight years), and Philippines (seven years). In others, such as Australia and the United Kingdom, both classifications coincide.

2.2 Propensity Score Estimations

In order to estimate (3), we built a yearly data set for 109 countries containing a set of variables that broadly define an economy ($\boldsymbol{X}$). The sources were the Penn World Table (PWT version 6.0) for GDP per capita and national accounts data; the *IFS* for international reserves, money, and credit markets data; and Reinhart and Rogoff (2004) for exchange rate regimes.[22]

The variables entered in the regression are the averages of the five years prior to the IT adoption for ITers. To check for robustness, for non-ITers we use either the average since 1990 up to 2004 or the five years previous to 1996 (for classification 1) or 1998 (for classification 2).[23] As described earlier, we tested for the balancing hypothesis and selected the most parsimonious specification.

In table 2 we show the variables whose coefficients were statistically significant in the four estimated models: from the PWT, investment to GDP, exports plus imports to GDP (namely, openness ratio), and the share of world GDP (GDP for a particular country to the sum of GDPs of the 109 countries in the database); from the IFS, the fiscal balance to GDP, inflation and its coefficient of variation (inflation volatility), and the money-to-GDP ratio; finally, the average number of years that a country was classified as *freely floating* by Reinhart and Rogoff (2004).

[22] We also considered social indicators from the World Bank and other sources for central bank staff and geographical controls. These variables were not significative in the regressions.

[23] These are the average adoption dates in each classification.

Table 1. Inflation Targeters and Dates of Adoption

	Corbo, Landarretche, and Schmidt-Hebbel (2002)	Fracasso, Genberg, and Wyplosz (2003)	Fraga, Goldfajn, and Minella (2003)	Levin, Natalucci, and Piger (2004)	Pétursson (2004)	Ball and Sheridan (2005)	Class. 1	Class. 2
Australia	1994	1994	1993		1993	1994	1994	1994
Brazil	1999	1999	1999	1999	1999		1999	1999
Canada	1991	1991	1991		1991	1992(94)	1991	1994
Chile	1991	1991	1991	1991	1990		1991	1999
Colombia	1999	1999	1999	1999	1999		1995	1999
Czech Republic	1998	1998	1998	1998	1998		1998	1998
Finland†	1993					1994	1993	1993
Hungary		2001	2001	2001	2001		2001	2001
Iceland	2001	2001	2001		2001		2001	2001
Israel	1992	1992	1992	1992	1992		1992	1997
Mexico	1999	1999	1999	1999	1999		1995	1999
New Zealand	1990	1988	1990		1990	1990(93)	1990	1991
Norway	2001	2001	2001	2001	2001		2001	2001
Peru	1994	2002	1994	2002	2002		1994	2002
Philippines		2002			2002		1995	2002
Poland	1998	1998	1998	1998	1998		1998	1998
South Africa	2000	2000	2000	2000	2000		2000	2000
South Korea	1998	1998	1998	1998	1998		1998	1998
Spain†	1995					1994(95)	1994	1995
Sweden	1993	1993	1993		1993	1995	1993	1995
Switzerland	2000	2000	2000	2000	2000		2000	2000
Thailand	2000	2000	2000	2000	2000		2000	2000
United Kingdom	1992	1992	1992		1992	1993	1992	1992

Note: Blank cells mean the authors did not provide a clear reference of the date of IT adoption.
† Finland and Spain abandoned inflation targeting and adopted the euro in 1999.

Table 2. Propensity Score Estimation, Logit Regressions

	(1)		(2)		(3)		(4)	
Classification for ITers	Class. 1		Class. 2		Class. 1		Class. 2	
Classification for Non-ITers	> 1990		> 1990		Class. 1		Class. 2	
Investment to GDP	0.337	(0.099)	0.250	(0.073)	0.402	(0.111)	0.282	(0.076)
Openness Ratio	−0.057	(0.012)	−0.042	(0.013)	−0.010	(0.027)	−0.065	(0.019)
Share of World GDP	−0.591	(0.199)	−0.342	(0.161)	−0.712	(0.313)	−0.437	(0.244)
Fiscal Balance to GDP	0.291	(0.166)	0.147	(0.103)	0.325	(0.150)	0.159	(0.120)
CPI Inflation	0.428	(0.133)	0.254	(0.099)	0.351	(0.126)	0.242	(0.097)
Inflation Volatility	−5.206	(1.926)	−3.599	(1.543)	−4.523	(1.957)	−2.929	(1.752)
Money to GDP	0.033	(0.015)	0.027	(0.013)	0.051	(0.021)	0.028	(0.015)
Exchange Rate Regime	−0.232	(0.079)	−0.154	(0.061)	−0.207	(0.079)	−0.141	(0.055)
Observations	100		100		100		100	
Pseudo R^2	0.6114		0.4704		0.6066		0.4940	
LR Stat, $\chi^2(8)$	65.95		50.74		65.43		53.28	
Common Support Region	[0.036, 0.998]		[0.037, 0.994]		[0.030, 0.993]		[0.015, 0.995]	
Non-ITers in Common Support	28		31		30		43	

Note: Figures in parentheses are robust standard errors.

Figure 1. Propensity Score Densities by IT Adoption Date

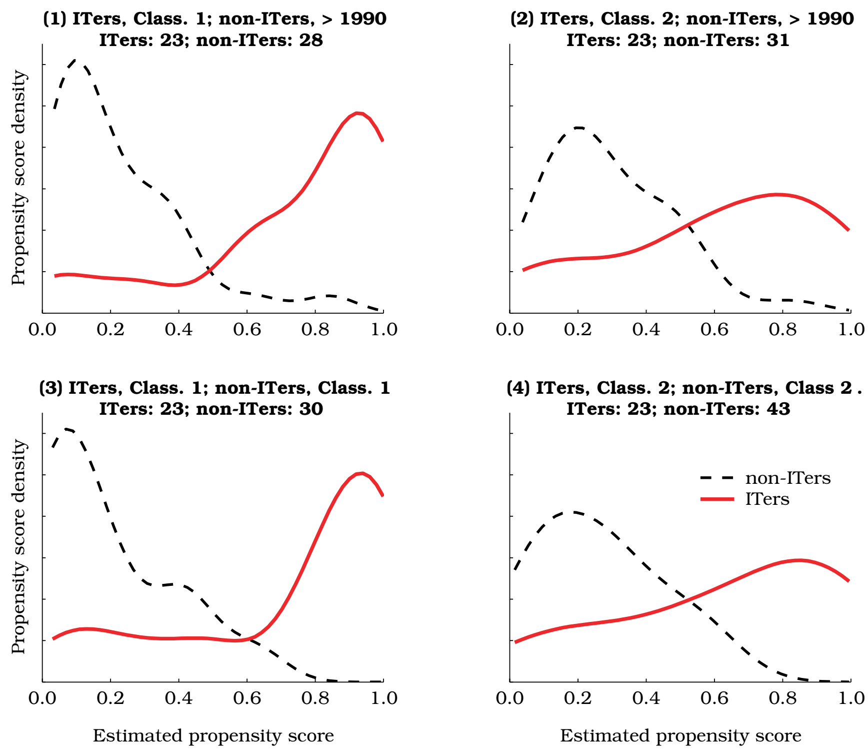

Figure 1 displays the density of the propensity score for ITers and non-ITers derived for each of the estimated models. It can be seen that the densities for model (1) are close to those of model (3); similarly, model (4) resembles model (2). For this reason, we will work with the first two specifications, where the differences between the propensities scores are driven by the alternative IT adoption dates, and not by variations in the control group.

2.3 Inflation Outcomes

A shortcoming of working with a wide control group is the low availability of data. Even though the consumer price index (CPI) time series are readily available for most of the countries, this is not true with some interesting variables. Such is the case for inflation expectations (from surveys) or forecast errors (from polls) that are directly influenced by IT adoption[24] or cross-sectional higher moments (skewness and kurtosis) of the CPI distribution.

Hence, the outcomes we use are quantities that can be extracted from conventional CPI data that broadly characterize inflation dynamics: level, variation, and persistence. We built a yearly data set from quarterly CPI information from the IMF's database (IFS), computed the counterfactuals, and estimated the IT effects as in (6) and (7).[25] For each year t the *level* of inflation is defined as the mean of the annualized quarterly inflation rates of years t and $t-1$. The same logic applies to the *standard deviation* of inflation, to measure volatility.

The interesting debate on measuring inflation *persistence*[26] can be summarized in the equation

$$\pi_t - \mu_t = \rho(\pi_{t-1} - \mu_{t-1}) + \sum_{\tau=1}^{p} \beta_\tau \Delta(\pi_{t-\tau} - \mu_{t-\tau}) + \epsilon_t \qquad (8)$$

that is a reparameterization of a simple AR(p) process for $(\pi_t - \mu_t)$, the deviation of inflation (π_t) from its mean (μ_t). A common practice is to set $\mu_t = \mu$ and estimate the parameter ρ, which equals the sum of all the autoregressive coefficients in the original AR(p) representation.[27] The closer ρ is to one, the more persistent the inflation.

However, Robalo Marques (2004) has pointed out that if the true process in (8) has a time-varying mean, imposing $\mu_t = \mu$ leads to misleading conclusions. Particularly, a series that quickly reverts to a

[24]See Johnson (2002) for an application to a sample of selected countries.

[25]As a baseline we consider the pretreatment period to be the average of the five years before the IT adoption (k in equation [6]), as we did in the propensity score estimations. We also tried different definitions, though the results were not sensitive to this assumption.

It is important to note that the number of years after IT (T_i in equation [7]) varies as IT adoption dates do. For classification 1 [2] there are $\sum_{i=1}^{N} T_i = 175\,[132]$ post-IT observations.

[26]See Robalo Marques (2004) for a survey. This author also shows that the approach followed here to measure persistence, even though it has some limitations, seems to be the most reliable among simple alternatives.

[27]It is well known that the OLS estimator of ρ is biased when $\rho \simeq 1$. An alternative (and popular) estimator, which is adopted here, is proposed in Andrews and Chen (1994).

time-varying mean may be estimated as highly persistent (ρ close to one) if it is assumed to revert to an imposed constant level. To control for this undesirable effect, he suggests estimating μ_t as a smooth trend of π_t. Considering this, we use two measures of inflation persistence: the estimated ρ with $\mu_t = \mu$ and with μ_t approximated by the HP filter.[28] To compute these quantities we use rolling windows with between ten and fifteen years of quarterly data.[29]

3. The Effects of Inflation Targeting

In table 3 we present the estimated average effects of IT for all ITers, for the group of industrialized countries as well as developing ones. We report effects on inflation dynamics according to our two alternative classifications of IT adoption. In the spirit of the mean-regression hypothesis of Ball and Sheridan (2005), we also include the results obtained by controlling for initial (pretreatment) conditions.[30]

The first key result is that IT has significantly reduced mean inflation in all the cases. In general, we find that the benefits of *soft* IT adoption are stronger than those of *fully fledged* IT adoption. This was expected due to high-inflation countries adopting IT to stabilize (the dates in classification 1). Also, the benefits on developing countries have been significantly stronger than those on industrialized ones, which confirms previous findings in Bernanke et al. (1999), Corbo, Landarretche, and Schmidt-Hebbel (2002), Neumann and von Hagen (2002), and Pétursson (2004). The results also suggest that regression to the mean is indeed an important phenomenon, since the effects of IT tend to be smaller once we control for initial conditions. However, by considering substantially wider treatment and control groups than the ones in Ball and Sheridan (2005), we find that there is no sufficient evidence to discard the benefits of IT: IT matters for mean inflation in both industrial and developing countries.

[28] We use a smoothing parameter of $\lambda = 1,600$. Different choices of λ do not qualitatively change the results.

[29] The lag length in (8), p, was selected to minimize the Schwarz criterion.

[30] To control for initial conditions as in Ball and Sheridan (2005), we compute the average treatment effects on the treated on a new variable $e_{i,t}$, which is obtained as the residual of the regression $Y_{i,t} - Y_{i,t'} = \alpha + \beta Y_{i,t'} + e_{i,t}$.

Table 3. Average Treatment Effect of IT on ITers

		All ITers		Industrialized Countries		Developing Countries	
Classification 1	Level	−4.802	(0.440)	−3.335	(0.627)	−6.320	(0.631)
Difference	Standard Deviation	−2.099	(0.323)	−1.546	(0.468)	−2.671	(0.452)
in means	Persistence ($\mu_t = \mu$)	0.027	(0.042)	0.031	(0.068)	0.024	(0.050)
	Persistence ($\mu_t =$ HP)	−0.028	(0.026)	−0.092	(0.023)	−0.039	(0.011)
Classification 2	Level	−2.863	(0.235)	−1.327	(0.334)	−5.382	(0.297)
Difference	Standard Deviation	−1.551	(0.318)	−1.103	(0.386)	−2.286	(0.557)
in means	Persistence ($\mu_t = \mu$)	0.027	(0.032)	0.003	(0.047)	0.066	(0.036)
	Persistence ($\mu_t =$ HP)	−0.016	(0.024)	−0.061	(0.018)	−0.058	(0.012)
Classification 1	Level	−3.874	(0.745)	−2.804	(0.868)	−4.907	(1.269)
Regression,	Standard Deviation	−1.863	(0.413)	−0.988	(0.568)	−2.708	(0.657)
controlling for	Persistence ($\mu_t = \mu$)	0.030	(0.039)	0.012	(0.057)	0.049	(0.058)
initial conditions	Persistence ($\mu_t =$ HP)	−0.015	(0.031)	−0.006	(0.022)	−0.023	(0.024)
Classification 2	Level	−2.621	(0.312)	−1.603	(0.421)	−3.242	(0.337)
Regression,	Standard Deviation	−1.798	(0.308)	−1.284	(0.383)	−2.112	(0.478)
controlling for	Persistence ($\mu_t = \mu$)	0.043	(0.023)	0.012	(0.035)	0.094	(0.035)
initial conditions	Persistence ($\mu_t =$ HP)	−0.047	(0.021)	−0.033	(0.016)	−0.055	(0.016)

Note: Figures in parentheses are bootstrapped standard errors (5,000 replications).

As mentioned in Faust and Henderson (2004), "Common wisdom and conventional models suggest that best-practice policy can be summarized in terms of two goals: first, get mean inflation right; second, get the variance of inflation right." Our finding regarding mean inflation supports the idea that IT in fact helps in achieving the first goal. What about the second goal? During the period of analysis, inflation has been falling worldwide, and together, the variance of inflation has been decreasing everywhere as well.[31] Our second finding precisely indicates that the observed fall in the variance of inflation has been particularly strong within ITers, such that the treatment effect has been that of a marked reduction in inflation volatility. The pattern of this effect across country groups and IT classifications is similar to the one found for the level of inflation. Neumann and von Hagen (2002) and Corbo, Landarretche, and Schmidt-Hebbel (2002) also provide evidence suggesting that IT has contributed to the fall in inflation volatility.[32]

What can we say about IT effects on inflation persistence? As mentioned, there is no straightforward theoretical prediction of the effects of IT on persistence. Adoption of IT can be linked to either lower or higher inflation persistence; it all hinges on two opposing effects: how fast central banks allow inflation to revert back to its mean after a shock and how price formation changes if expectations become more anchored. Studies like Levin, Natalucci, and Piger (2004) show that persistence is lower in ITers than that in non-ITers, whereas Ball and Sheridan (2005) show there is no evidence that ITers achieve lower inflation persistence.[33]

We find that the results depend on the measure of persistence (ρ) used. If we consider a constant unconditional mean in the inflation process ($\mu_t = \mu$), we find that IT increases persistence, though the estimates are not statistically significant and different from zero. Contrarily, if we allow for a time-varying mean inflation ($\mu_t = \text{HP}$),

[31] See Pétursson (2004).

[32] Johnson (2002) and Ball and Sheridan (2005) suggest that IT increases inflation uncertainty. The finding in Johnson (2002) in fact refers to volatility of expected inflation from surveys, a variable related to observed inflation volatility but with dynamics of its own.

[33] Time-series studies on persistence for industrial countries—like those of Benati (2004), Levin and Piger (2004), or Robalo Marques (2004)—point to the conclusion that high inflation persistence is not a robust feature of inflation processes in the euro area or the United States.

we find that IT does reduce the persistence parameter. Interestingly, some sort of mean regression is present under classification 1 (soft IT): once we control for the initial persistence, the fall in ρ disappears. However, under classification 2 (fully fledged IT) the fall in ρ is significant even after controlling for mean regression (which seems to exist in industrialized economies).

This last effect, although different from zero, is at most modest. The half life of a shock to inflation is, roughly speaking, $\tau \approx -\ln(2)/\ln(\rho)$.[34] The change in ρ implied by our results varies around -0.04; hence, considering an initial $\rho = 0.85$,[35] the change in τ is just one quarter. All in all, the evidence on the effect of IT on inflation persistence, if any, is not as categorical as the one associated with the reduction in mean and volatility.

4. Concluding Remarks

The increasing popularity of IT as a framework for conducting monetary policy calls for the evaluation of its benefits in comparison to alternative schemes. In this study we have combined data of IT adoption and inflation dynamics with program evaluation techniques to assess the dimensions in which IT is a beneficial regime. Our central findings support the idea that the adoption of IT, either in its soft or explicit form, delivers the theoretically promised outcomes: low mean inflation (around a fixed target or within a target range) and low inflation volatility.

We also find that IT has reduced the persistence of inflation in developing countries. Given that IT is understood to be flexible, the reduction in persistence is likely to be the effect of the anchoring of expectations to a defined nominal level. Nevertheless, the small magnitude of the reduction is such that it prevents us from categorically concluding in favor of IT in this particular dimension of the inflation dynamics. In the future, it would be useful to contrast our results with alternative measures of persistence. Also, a promising area for further research is to formalize the theoretical link between IT, inflation persistence, and long-run expectations (credibility), which can guide subsequent empirical efforts.

[34] This formula is exact if the estimated model is an AR(1).

[35] This is a generous value. The sample mean of all our computed ρ after detrending is just below 0.40.

The interpretation we gave to IT adoption, that of a "natural experiment," allowed us to use powerful evaluation tools normally applied in microeconometrics. Yet, it is important to keep in mind that the identifying assumption in a macroeconomic context like ours might be stronger. We also reckon that the study of the response of other macroeconomic variables (for instance, the business cycles and interest rates) to IT is essential in order to have a complete appraisal of the effects of the IT regime. Hence, future research can explore further, within the IT adoption evaluation, the advantages of these techniques on a wider variety of macro indicators.

References

Andrews, Donald W.K., and Hong-Yuan Chen. 1994. "Approximately Median-Unbiased Estimation of Autoregressive Models." *Journal of Business and Economic Statistics* 12 (2): 187–204.

Ball, Laurence, and Niamh Sheridan. 2005. "Does Inflation Targeting Matter?" In *The Inflation Targeting Debate*, NBER Studies in Business Cycles, Vol. 32, ed. Ben S. Bernanke and Michael Woodford. Chicago: University of Chicago Press.

Benati, Luca. 2004. "International Evidence on Inflation Persistence." Mimeo, Bank of England.

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

Corbo, Vittorio, Oscar Landarretche, and Klaus Schmidt-Hebbel. 2002. "Does Inflation Targeting Make a Difference?" In *Inflation Targeting: Design, Performance, Challenges*, ed. Norman Loayza and Raimundo Soto, 221–69. Central Bank of Chile.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84 (1): 151–61.

Faust, Jon, and Dale W. Henderson. 2004. "Is Inflation Targeting Best-Practice Monetary Policy?" International Finance Discussion Papers 807, Board of Governors of the Federal Reserve System.

Fracasso, Andrea, Hans Genberg, and Charles Wyplosz. 2003. "How Do Central Banks Write? An Evaluation of Inflation Reports by Inflation-Targeting Central Banks." Geneva Reports on the World Economy Special Report 2. International Center for Monetary and Banking Studies, Centre for Economic Policy Research (CEPR) and Norges Bank.

Fraga, Arminio, Ilan Goldfajn, and André Minella. 2003. "Inflation Targeting in Emerging Market Economics." NBER Working Paper No. 10019.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing selection bias using experimental data." *Econometrica* 66 (5): 1017–98.

Heckman, James J., Hidehiko Ichimura, and Petra Todd. 1997. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *Review of Economic Studies* 64 (4): 605–54.

———. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65 (2): 261–94.

Johnson, David R. 2002. "The effect of inflation targeting on the behavior of expected inflation: evidence from an 11 country panel." *Journal of Monetary Economics* 49 (8): 1521–38.

Levin, Andrew T., Fabio M. Natalucci, and Jeremy M. Piger 2004. "The Macroeconomic Effects of Inflation Targeting." *Review* (Federal Reserve Bank of St. Louis) 86 (4): 51–80.

Levin, Andrew, and Jeremy Piger. 2004. "Is Inflation Persistence Intrinsic in Industrial Economies?" Working Paper No. 334, European Central Bank.

Mishkin, Frederic S., and Klaus Schmidt-Hebbel. 2002. "One decade of inflation targeting in the world: What do we know and what do we need to know?" In *Inflation Targeting: Design, Performance, Challenges*, ed. Norman Loayza and Raimundo Soto, 171–219. Central Bank of Chile.

Neumann, Manfred J.M., and Jürgen von Hagen. 2002. "Does Inflation Targeting Matter?" *Review* (Federal Reserve Bank of St. Louis) 84 (4): 127–48.

Pétursson, Thórarinn G. 2004. "The effects of inflation targeting on macroeconomic performance." Working Paper 23, Central Bank of Iceland.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2004. "The Modern History of Exchange Rate Arrangements: A Reinterpretation." *Quarterly Journal of Economics* 119 (1): 1–48.

Robalo Marques, Carlos. 2004. "Inflation Persistence: Facts or Artifacts?" Working Paper No. 371, European Central Bank.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Casual Effects." *Biometrika* 70 (1): 41–55.

Smith, Jeffrey A., and Petra Todd. 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics* 125 (2): 305–53.

Svensson, Lars E.O. 1999. "Inflation Targeting: Some Extensions." *Scandinavian Journal of Economics* 101 (3): 337–61.

———. 2000. "Open-economy inflation targeting." *Journal of International Economics* 50:155–83.

Truman, Edwin M. 2003. *Inflation Targeting in the World Economy.* Washington, DC: Institute for International Economics.