Inflation Targeting and Inflation Behavior: A Successful Story?*

Marco Vega†
London School of Economics &
Central Bank of Peru
Jr. Miroquesada 441, Lima 1, Peru
m.a.vega@lse.ac.uk

Diego Winkelried
St Edmund's College,
University of Cambridge
CB3 0BN, Cambridge, UK
dw295@cam.ac.uk

February 15, 2005

ABSTRACT

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 determine suitable counterfactuals to the actual inflation targeters. With this approach we find 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. We also find that the effect of IT in the persistence of inflation is rather weak and not as categorical as the one associated with the mean and volatility of inflation.

Keywords : Inflation targeting, matching methods.
JEL Codes : C50, E42, E52.

*We are grateful to seminar participants at the Central Bank of Peru for their valuable comments. The opinions hereby are those of the authors, not necessarily of the institutions.

†Corresponding author: ++ 511 613 2042.

1 Introduction

The adoption of an inflation targeting regime (IT from now on) as a monetary policy framework since the early 90's by an increasing number of central banks has spurred research on the benefits of such 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, 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 asses 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 pre-adoption 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 interesting 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 had 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. The reduction of endogenous uncertainty arises from increasing the signal-to-noise ratios of central bank policy

[1] To date, twenty one countries follow an explicit IT framework and there are some countries that are considering its adoption. See for example Truman (2003) and Pétursson (2004).
[2] See for example Svensson (2000).
[3] See for Bernanke et al. (1999), Mishkin and Schmidt-Hebbel (2002) and Corbo et al. (2002).
[4] See Neumann and Von Hagen (2002) for a recent empirical survey.

instruments. Modern monetary policy practice, not exclusively IT, hinges precisely on making monetary policy more predictable. 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 rip some benefits of flexibility, or the nominal anchor is lost and the inflation falls to the expectation trap.

Hence, either more or less persistence can result from IT adoption. Taking aside any other shock and focusing just on monetary policy and credibility, a reduction in persistence may easily be the result for emerging market economies being stricter ITers looking to strengthen credibility. More persistence can result from successful flexible ITers or unsuccessful ITers not gaining credibility. Once more, what would the empirical evaluation of IT over persistence tell us about the adopting ITers?

In recent years, a growing literature has provided insights on the empirical assessment of IT. Corbo et al. (2002), for instance, compare policies and outcomes in fully-fledged IT countries to two groups, potential ITers and non-ITers. Among other things, 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. Even though the richness in the measures of macroeconomic outcomes to evaluate and the comparison to a control group, Corbo et al. (2002) is yet an evaluation under the absolute criteria.

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. On the contrary, the paper finds that IT has not reduced absolute average forecast errors in targeting countries relative to those in non-targeting countries. However, ITers did avoid even larger forecast errors than 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

[5] See Svensson (1999).

inflation and long-run as well as short-run interest rates to a negative supply shock (increase in world oil prices[6]). They find that the effect of IT is not significantly different from zero for average inflation, but significant for both short and long interest rates, meaning that the gain in inflation credibility has been superior in IT countries.

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. His overall finding is that IT has indeed been beneficial to reduce average inflation, persistence and the variability of inflation[7]. However, the technique offered by this paper, does not tackle the fundamental question of relative performance. Its contribution hinges in giving a clear account for the evidence of the absolute benefits of IT and corroborates previous findings on this line.

Levin et al. (2004) study inflation persistence using five industrial ITers which are compared to seven industrial non-ITers. They perform univariate regressions on inflation for each country and find that inflation persistence is estimated to be quite low within ITers whereas the unit root hypothesis cannot be rejected for non-ITers. On the other hand, Levin and Piger (2004) in a similar empirical framework with twelve industrial countries, allow for structural breaks and find that inflation in general exhibits low persistence[8]. Additionally, they 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.

Finally, Ball and Sheridan (2005) provide evidence on the irrelevance of IT. They look at seven OECD countries that adopted IT in the early 90's and thirteen countries that did not. They claim that ITers that reduced higher-than-average inflation rates towards 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 find that there is no evidence that IT improves 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. This is precisely the reason why this matter of comparability will not be overlooked here.

[6] This type of shock creates a dilemma because they are inflationary and at the same time generate a downturn of economic activity. They choose two episodes: 1978-1979 and 1998-1999.

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

[8] These results confirm those of Benati (2004) that studies inflation dynamics in twenty OECD countries and Emery (1994) that analyzes the USA postwar inflation process.

Following Neumann and Von Hagen (2002), 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 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 and the *widest possible control group* of non-ITers (86 countries) using 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, given that IT adoption can be interpreted as a natural experiment, we aim to reestablish the conditions of a randomized experiment and represent IT adoption as 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 2 we briefly describe the propensity score and matching techniques for evaluation; in section 3 we discuss some empirical issues regarding the robustness of our results and present the inflation outcomes to be evaluated; in section 4 we show our main findings and section 5 concludes and provides some avenues for future research.

2 Methodology

As mentioned, we use microeconometric techniques usually applied in non-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.

2.1 The fundamental problem

Let D be a binary indicator that equals unity if a country has adopted IT and zero otherwise. Also, let Y_t^1 denote the value of 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

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

It is clear from (1) that we face an identification problem since $E[Y_t^0|\boldsymbol{X}, D = 1]$ 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

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

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 (i.e. most developed countries having less volatile inflation rates). A common approach to estimate the expectation $E[(Y_t^0 - Y_{t'}^0)|\boldsymbol{X}, D = 1]$ is to replace it with the observable average outcome in the untreated state $E[(Y_t^0 - Y_{t'}^0)|\boldsymbol{X}, D = 0]$ and, hence, consider the statistic

$$\hat{\xi}_{dd} = E\left[(Y_t^1 - Y_{t'}^0)|\boldsymbol{X}, D = 1\right] - E\left[(Y_t^0 - Y_{t'}^0)|\boldsymbol{X}, D = 0\right] \tag{3}$$

which is known as the *difference-in-difference* estimator[9].

However, normally $E[Y_t^0|\boldsymbol{X}, D = 1] \neq E[Y_t^0|\boldsymbol{X}, D = 0]$, so (3) will render biased estimates of ξ from two sources[10]. 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 the $\boldsymbol{X}$ between the treated and the control groups, which is usual in non-randomized samples (like a dataset of countries). Fortunately, matching methods deal with these shortcomings.

2.2 Matching methods

The idea behind matching techniques is 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, to reestablish the conditions of a randomized experiment (that is, random assignment of the $\boldsymbol{X}$) when no such data are available. Under these circumstances, the difference between the outcome of the treated and that of a matched counterfactual can be attributed to the treatment effect.

2.2.1 The propensity score

Usually, determining along which dimension to match the countries or what type of weighting scheme 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,

$$p(\boldsymbol{X}) = E\left[D|\boldsymbol{X}\right] = \Pr\left(D = 1|\boldsymbol{X}\right) \tag{4}$$

[9] This estimator encompasses a widely-used practice: under homogeneous treatment effects across the population, i.e. $E\left[(Y_t^1 - Y_t^0)|\boldsymbol{X}, D = 1\right] = a$, $\hat{\xi}_{dd}$ is equivalent to the least square estimate of a in the panel equation $Y_{i,t} = aD_{i,t} + \boldsymbol{X}_{i,t}\boldsymbol{b} + \boldsymbol{c}_i + \boldsymbol{d}_t + \epsilon_{i,t}$. See Johnson (2002) and Ball and Sheridan (2005).

[10] See Heckman et al. (1998a).

and should satisfy the *balancing hypothesis*, which states that observations with the same propensity score must have the same distribution of observable and unobservable characteristics independently of the treatment status, formally $D \perp \boldsymbol{X}|p(\boldsymbol{X})$[11]. Hence, equation (1) can be rewritten as

$$\hat{\xi} = E\left[(Y_t^1 - Y_{t'}^0)|p(\boldsymbol{X}), D = 1\right] - E\left[(Y_t^0 - Y_{t'}^0)|p(\boldsymbol{X}), D = 1\right] \tag{5}$$

The first source of bias (non-comparability among ITers and non-ITers) can be eliminated by only considering countries within the *common support*, the interval on the real line where both distributions $\{p(\boldsymbol{X})|D = 1\}$ and $\{p(\boldsymbol{X})|D = 0\}$ have positive densities. The second source of bias (difference in the distribution of the observable variables) is eliminated by reweighing the non-ITers observations. This is the very goal of matching methods: conditional on $\boldsymbol{X}$, to equalize the counterfactual outcome distribution of the non-ITers with the observed outcome distribution of the non-ITers.

Estimating the propensity score is straightforward, as any probabilistic model suits (4). For instance, we can adopt the parametric form $\Pr(D_i = 1|\boldsymbol{X}_i) = F(h(\boldsymbol{X}_i))$ where $F(.)$ is the logistic cumulative distribution (i.e. a logit). However, two points are to be handle 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[12]. Second, the model selection, i.e. the form of $h(\boldsymbol{X}_i)$, can be seen as a way of testing the balancing hypothesis. Dehejia and Wahba (2002) suggest using a polynomial according to the following steps:

- Start with a parsimonious logit specification (i.e. $h(\boldsymbol{X}_i)$ linear)
- 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 $(0 - 0.2, \ldots, 0.8 - 1)$.
- For each interval, test if the averages of $\boldsymbol{X}$ of treated and control units do not differ. If covariates are balanced between these groups for all strata, the specification satisfies the balancing hypothesis[13]. If the test fails in one interval, divide it into smaller strata and reevaluate.

[11] Matching methods rely on the fundamental assumption of *conditional independence* between outcomes and the treatment, i.e. $\{Y^1, Y^0\} \perp D|\boldsymbol{X}$. This assumption states that given $\boldsymbol{X}$, the non-treated outcomes are what the treated outcomes would have been had they not been treated (therefore, for each treated we could in principle find a control observation with the same $\boldsymbol{X}$). Rosenbaum and Rubin (1983) and Rosenbaum and Rubin (1984) show that if the balancing hypothesis holds, conditioning on $p(\boldsymbol{X})$ instead preserves this important condition (i.e. $\{Y^1, Y^0\} \perp D|p(\boldsymbol{X})$). In practice, we require a weaker and *testable* condition to identify the treatment effect: conditional mean independence, $E[Y^0|\boldsymbol{X}, D = 1] = E[Y^0|\boldsymbol{X}, D = 0]$.

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

[13] Actually, the weaker version of mean conditional independence. See footnote 11.

- 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.

2.2.2 The matched difference-in-difference estimator

Given the propensity score, there are various methods available for finding a counterfactual for ITer i[14]. Following Heckman et al. (1997) and Heckman et al. (1998a), 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-ITers countries whose propensity scores are over the region of the common support. The counterfactual of the outcome $Y_{i,t}^0$ is

$$\tilde{Y}_{i,t}^0 = \frac{\sum_{j\in\mathcal{C}} K(\frac{p_j - p_i}{h}) Y_{j,t}^0}{\sum_{j\in\mathcal{C}} K(\frac{p_j - p_i}{h})} \tag{6}$$

where $K(.)$ is a kernel function (with bandwidth parameter h) that weights the outcome of country j inversely proportional to the distance between its propensity score value (p_j) and the one of the non-ITer i (p_i).

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

$$\hat{\xi}_{i,t} = \left(Y_{i,t}^1 - \frac{1}{k-1} \sum_{\tau=1}^{k-1} Y_{i,\tau}^0 \right) - \left(\tilde{Y}_{i,t}^0 - \frac{1}{k-1} \sum_{\tau=1}^{k-1} \tilde{Y}_{i,\tau}^0 \right) \tag{7}$$

where the pre-treatment 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[15]. The estimator (7) 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).

3 Empirical issues

Before presenting 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.

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

[15] Heckman et al. (1998a) and Smith and Todd (2005) suggest using a weighted average of the pre-treatment observations instead of a sole observation to control for possible outliers or trend effects. In (7) we have used a simple average (equal weights).

3.1 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 and for the sake of robustness, we use two possible adoption dates for each country[16]. 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 (forecast)[17].

Our approach contrasts 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[18]; 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 run exchange rate regimes not compatible with fully-fledged IT until 1999. For Peru, authors such as Corbo et al. (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 et al. (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), 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 and since then it has remained unchanged.

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 for December 1995 even though the Riksbank announced its shift to IT during 1993. For Canada, the first target range was announced in 1991. In December 1993, a range of 1 to 3 percent was established for 1994 onwards.

In Table 1 we compare adoption dates among five different studies and provide our two possible adoption dates. Column "Class. 1" refers to the soft IT adoption dates while "Class. 2" accounts for fully-fledged IT adoption. In 6 cases we have more than a three-year difference

[16] We also perform estimations with different samples for the non-ITers observations. See the Appendix.

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

[18] See Fraga et al. (2003) for a comprehensive survey of IT in developing countries.

Table 1: **Inflation targeters and dates of adoption** [a]

	Corbo et al. (2002)	Fracasso et al. (2003)	Fraga et al. (2003)	Levin et al. (2004)	Pétursson (2004)	Ball & Sheridan (2004)	Class. 1 [b]	Class. 2 [b]
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 [c]	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 [c]	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

[a] Blank cells mean the authors did not provide a clear reference of the date of IT adoption.
[b] Our classifications come from each central bank's webpage. See the main text for details.
[c] Finland and Spain abandoned inflation targeting and adopted the Euro in 1999.

between both dates: Chile (8 years), Colombia (4 years), Israel (5 years), Mexico (4 years), Peru (8 years) and Philippines (7 years). In others, such as Australia and the UK, both classifications coincide.

3.2 Inflation outcomes

One shortcoming of working with a wide control group is 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 forecasts errors (from polls) that are directly influenced by IT adoption[19] 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 dataset from quarterly CPI information from the IMF's database (IFS), compute the counterfactuals and find the treatment effect on a country basis[20]. The average over time and ITers from $\hat{\xi}_{i,t}$ in equation (7) is the estimate of interest[21]. 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.

The interesting debate on measuring inflation *persistence*[22] can be summarized in the equation

$$\pi_t - \mu_t = \rho(\pi_{t-1} - \mu_{t-1}) + \sum_{j=1}^{p} \beta_j \Delta(\pi_{t-j} - \mu_{t-j}) + \epsilon_t \tag{8}$$

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[23]. The closer is ρ 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 ρ will capture two effects: persistence and *mean-reversion*, so it is possible to obtain a ρ close to one with a series that is not persistent at all. To control for this undesirable effect, he suggests, within a univariate framework, to estimate 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[24]. To compute these quantities we use rolling windows with between 10 and 15 years of quarterly data[25].

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

[20] As a baseline we consider the pre-treatment period to be 5 years before the IT adoption (k in equation (7)). We also tried different definitions, though the results were not sensitive to this assumption.

[21] It is important to note that the number of years after IT varies as IT adoption dates do. For Classification 1 [2] there are 175 [132] post-IT observations.

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

[23] It is well known that the OLS estimator of ρ is biased when $\rho \simeq 1$. An alternative (and popular) estimator, that is adopted here, is proposed in Andrews and Chen (2004).

[24] We use a smoothing parameter of $\lambda = 1600$. Different choices of λ do not qualitatively change the results.

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

4 Results

In Table 2 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 (pre-treatment) conditions[26].

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-fledge* 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 et al. (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 a 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 alike.

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 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[27]. Our second finding precisely points 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 variance. 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 et al. (2002) also provide evidence suggesting that IT has contributed to the fall in inflation volatility[28].

What can we say about IT effects on inflation persistence? As mentioned, there is no a 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:

[26] That is, using the notation of section 2, the first group of estimates are the treatment effects on $Y_{i,t} - Y_{i,t'}$ whereas the second are the treatment effects on the residuals of the regression $Y_{i,t} - Y_{i,t'} = \alpha + \beta Y_{i,t'} + e_{i,t}$.

[27] See Pétursson (2004).

[28] 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 a dynamics of its own.

Table 2: **Average treatment effect of Inflation Targeting** [a]

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

[a] Figures in parenthesis are bootstrapped standard errors (5000 replications).

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 et al. (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[29].

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.

[29] Time series studies on persistence for industrial countries like 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.

Contrary, if we allow for a time varying mean inflation (μ_t = HP) 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 seem 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, $\tau \approx -\ln(2)/\ln(\rho)$[30]. The changes in ρ implied by our results varies around –0.04; hence, considering an initial $\rho = 0.85$[31] 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 mean and volatility reduction.

5 Concluding Remarks

The increasing popularity of IT as a framework for conducting monetary policy claims 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 to conclude 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.

The interpretation we gave to IT adoption, that of a natural experiment, allowed us to use powerful evaluation tools normally applied in microeconometrics, where the odds to identify policy effects are by far higher than in macroeconomics. 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 if 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.

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

[31] This is a generous value. The sample mean of all our computed ρ after de-trending is just below 0.50.

References

Andrews, D. W. K. and H. Y. Chen (1994), Approximately median-unbiased estimation of autoregressive models, *Journal of Business & Economic Statistics*, 12(2), 187 - 204.

Ball, L. and N. Sheridan (2005), Does inflation targeting matter?, in B. S. Bernanke and M. Woodford (eds.), *The Inflation Targeting Debate*, The University of Chicago Press, forthcoming.

Benati, L. (2004), International Evidence on Inflation Persistence, Bank of England, mimeo.

Bernanke, B. S., T. Laubach, F. Mishkin and A. Posen (1999), *Infation Targeting: Lessons from the International Experience*, Princeton University Press.

Corbo, V., O. Landarretche and K. Schmidt-Hebbel (2002), Does inflation targeting make a difference?, in N. Loayza and R. Soto (eds.), *Inflation Targeting: Design, Performance, Challenges*, Central Bank of Chile, 221 - 269.

Dehejia, R. H. and S. Wahba (2002), Propensity score-matching methods for nonexperimental causal studies, *Review of Economics and Statistics*, 84(1), 151 - 161.

Emery, K. (1994), Inflation persistence and Fisher effects: evidence of a regime change, *Journal of Economic and Business*, 46, 141 - 152.

Faust, J. and D. Henderson (2004), Is inflation best-practice monetary policy?, International Finance Discussion Papers 807, Board of Governors of the Federal Reserve System.

Fraga, A., I. Goldfajn and A. Minella (2003), Inflation Targeting in Emerging Market Economies, NBER Working Paper 10019.

Heckman, J., H. Ichimura and P. Todd (1997), Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies*, 64(4), 605 - 654.

Heckman, J., H. Ichimura and P. Todd (1998a), Matching as an econometric evaluation estimator, *Review of Economic Studies*, 65(2), 261 - 294.

Heckman, J., H. Ichimura, J. Smith and P. Todd (1998b), Characterizing selection bias using experimental data, *Econometrica*, 66(5), 1017 - 1098.

Johnson, D. 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 - 1538.

Levin A., F. Natalucci and J. Piger (2004), The macroeconomic effects of inflation targeting, *Review*, Federal Reserve Bank of St. Louis, 86(4), 51 - 80.

Levin A. and J. Piger (2004), Is inflation persistence intrinsic in industrial economies?, European Central Bank, Working paper 334.

Mishkin, F. S. and K. Schmidt-Hebbel (2002), One decade of inflation targeting in the world: What do we know and what do we need to know?, in N. Loayza and R. Soto (eds.), *Inflation Targeting: Design, Performance, Challenges*, Central Bank of Chile, 171 - 219.

Neumann, M. J. M. and J. von Hagen (2002), Does inflation targeting matter?, *Review*, Federal Reserve Bank of St. Louis, 84(4), 127 - 148.

Pivetta, F. and R. Reis (2003), The persistence of inflation in the United States, mimeo, Harvard University.

Pétursson, T. G. (2004), The effects of inflation targeting on macroeconomic performance, Central Bank of Iceland, Working Paper 23.

Reinhart, C. M. and K. S. Rogoff (2004), The modern history of exchange rate arragemnets: A reinterpretation, *Quarterly Journal of Economics*, 119(1), 1 - 48.

Robalo Marques, C. (2004), Inflation persistence: Facts or artifacts?, European Central Bank, Working paper 371.

Rosenbaum, P. R. and D. B. Rubin (1983), The central role of the propensity socre in observational studies for casual effects, *Biometrika*, 70(1), 41 - 55.

Rosenbaum, P. R. and D. B. Rubin (1984), Reducing bias in observational studies using subclassifications on the propensity score, *Journal of the American Statistical Association*, 79, No. 387, 516 - 524.

Smith, J. and P. Todd (2005), Does matching address Lalonde's critique of nonexperimental estimators?, *Journal of Econometrics*, 125(2), 305 - 353.

Svensson, L. E. O. (1999), Inflation targeting: Some extensions, *Scandinavian Journal of Economics*, 101(3), 337 - 361.

Svensson, L. E. O. (2000), Open-economy inflation targeting, *Journal of International Economics*, 50, 155 - 183.

Truman, E. M. (2003) *Inflation Targeting in the World Economy*, Institute for International Economics, Washington.

A Appendix: Propensity score estimations

We present some details on the propensity score estimations under various definitions of IT adoption dates. It is important to recall that the role of the propensity score is to reduce the dimensionality of the matching, it does not necessarily convey a behavioral interpretation. Indeed, the logit regressions below 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 are useful to define the profile of a particular economy (importantly, relatively to others), are not theoretically included in the central bank's decision to change the monetary policy regime[32].

We built a yearly dataset for 109 countries containing a set of variables that broadly define an economy. 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, Reinhart and Rogoff (2004) for exchange rate regime, the World Bank for social indicators and other sources for central bank staff and geographical controls.

The variables entered in the regression are the averages of the five years previous 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 5 years previous to 1996 (for Classification 1) or 1998 (for Classification 2)[33]. As described in the text, we tested for the balancing hypothesis and selected the most parsimonious specification.

Table 3: **Propensity score estimation, logit regressions** [(a)]

	(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	

[(a)] Figures in parenthesis are robust standard errors.

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

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

In Table 3 above 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).

In Figure 1 we present 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 (2). For this reason, we work with the first two specifications in the text, where the differences between the propensities scores are driven by the alternative IT adoptions dates, and not by variations in the control group.

Figure 1: **Propensity score densities by IT adoption date**