MPRA
Munich Personal RePEc Archive

Characterizing the Anchoring Effects of Official Forecasts on Private Expectations

Barrera, Carlos

Banco Central de Reserva del Perú

19 August 2022

Online at https://mpra.ub.uni-muenchen.de/114258/
MPRA Paper No. 114258, posted 31 Aug 2022 13:27 UTC

CHARACTERIZING THE ANCHORING EFFECTS OF OFFICIAL FORECASTS ON PRIVATE EXPECTATIONS

Carlos R. Barrera-Chaupis*

Research Division, BCRP

August 31, 2022

Abstract

The paper proposes a method for simultaneously estimating the treatment effects of a change in a policy variable on a numerable set of interrelated outcome variables (different moments from the same probability density function). Firstly, it defines a non-Gaussian probability density function as the outcome variable. Secondly, it uses a *functional regression* to explain the density in terms of a set of scalar variables. From both the observed and the fitted probability density functions, two sets of interrelated moments are then obtained by simulation. Finally, a set of difference-in-difference estimators can be defined from the available pairs of moments in the sample. A stylized application provides a 29-moment characterization of the direct treatment effects of the Peruvian Central Bank's forecasts on two sequences of Peruvian firms' probability densities of expectations (for inflation $-\pi-$ and real growth $-g-$) during 2004-2015.

Keywords: statistical simulation methods, treatment effect models, central bank, forecasting, coordination.
JEL Classification JEL: C15, C30, E37, E47, E58, G14.

*The author would like to thank suggestions received from: Juan Carlos Aquino, Marco Vega, and seminar participants at the *Banco Central de Reserva del Perú* (BCRP).

I Introduction

The paper proposes a method for simultaneously estimating a numerable set of treatment effects (e.g. after a change in a policy variable) associated to the corresponding set of interrelated moments. Instead of using the temporal sequence of any specific moment (estimated from a sequence of large cross-sections), the paper uses the temporal sequence of probability density functions (estimated from such a sequence of large cross-sections). Therefore, the paper focuses on a single outcome variable, a general probability density function (not necessarily Gaussian), instead of focusing on many scalar outcome variables (one moment at a time).

The proposal's key ingredient is a *functional-regression* stage allowing to control for many scalar confounding explanatory variables. This regression substitutes a set of numerable (possibly non-linear) regressions, each explaining one scalar outcome variable. Then, a *simulation* stage that converts our useful outcome variable, the probability density function, into a numerable set of interrelated outcome variables (a set of moments obtained by simulation from the same probability density function).

The proposal is conceived to fully characterize the anchoring effects of a benevolent central bank' forecasts/announcements on private expectations (firms' or households') whenever private expectations consist on a temporal sequence of large cross-sections from which a temporal sequence of probability density functions can be obtained by non-parametric methods. As a byproduct, the simulation stage solves a problem inherent in the functional-regression stage, that the functional coefficients resulting from any functional regression have reduced interpretability.

The availability of such a sequence of large cross-sections (*big data*) may not be the only justification for this inquiry. Characterizations in the anchoring expectations literature usually consider at most two moments of those cross-sections of expectations (non-robust mean and dispersion) under the unwarranted assumption of Gaussianity.[1]

Figure 1: Jarque-Bera p-values (null: Gaussianity)

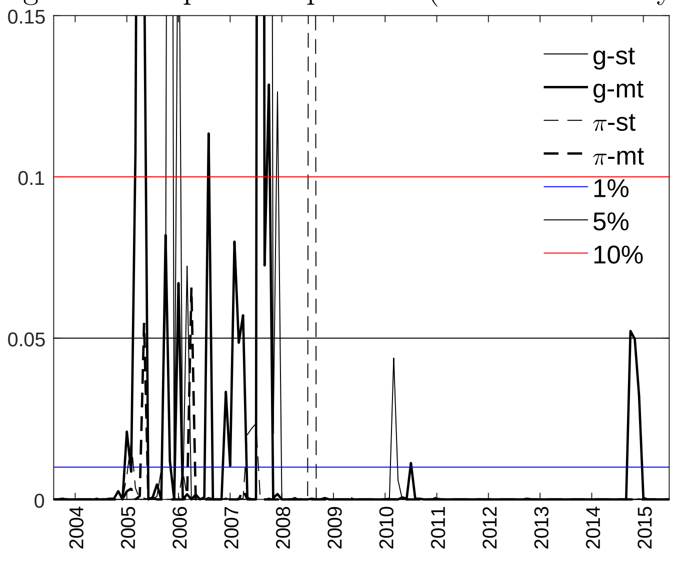


Figure 1 shows four temporal sequences of Jarque-Bera tests' p-values for the following monthly sequences of Peruvian firms' cross-sections of expectations during 2004-2015:

[1]The literature has usually provided results considering the mean of expectations without justifying their tools as useful enough (see [Blinder *et al.* (2008), Dräger *et al.* (2016), Filacek & Saxa (2012), Gürkaynak *et al.* (2010), Hattori *et al.* (2016), Kozicki & Tinsley (2005), Kumar *et al.* (2015), Neuenkirch (2013), Pereira da Silva (2016), Pedersen (2015), Trabelsi (2016)]). [Filacek & Saxa (2012)] and then [Barrera (2018)] use both dispersion and distance, which actually belong to different but related densities.

short-term (st) and medium-term (mt) real growth (g) expectations as well as st and mt inflation (π) expectations. The cross-sections hardly comply with Gaussianity!

For non-Gaussian data, an improved characterization of the anchoring effects of Central Bank' forecasts is not only possible from the paper's proposal, but needed out of the box. First, a benevolent central bank should care about the impact of its forecast (policy variable) on mainly robust versions of usual moments like the mean and the dispersion of firms' expectations. Second, a benevolent central bank should care about the impact of its inflation forecast on the probability mass of being in the target range, especially under the framework of inflation targeting. Interestingly, other moments can enhance the consistency of the aforementioned ones and thus should be included in a comprehensive set of moments to be simulated from the probability distribution of firms' forecasts.

To illustrate the usefulness of the proposal, a stylized application provides a moment characterization of the 'direct effects' of the Peruvian Central Bank's forecasts on two sequences of Peruvian firms' probability densities of expectations (of π and g) during 2004-2015. Main findings are: (i) Short-term π forecasts generates an *on-impact* increase in the probability that these expectations are in the target range of [1% 3%]. Short-term g forecasts generates an *on-impact* increase in the probability that these expectations are in the range of [4% 7%], but a *one-month-later* decrease in this probability. (ii) Medium-term π forecasts generates no significant changes in the probability that these expectations are in the target range of [1% 3%]. Medium-term g forecasts generates an *on-impact* decrease in the probability that these expectations are in the range of [4% 7%], but a *one-month-later* increase in this probability.[2]

Section II discusses the methodological issues associated to the functional regression models leading to the new complete-characterization tests. Section III describes the stylized application in terms of the Peruvian data (i.e., the probability density functions as the outcome variable, the central bank's forecasts as the treatment variable, as well as the control/explanatory variables) and the estimation results. Section IV concludes.

II Methodology

The proposal of this paper is closely tied to the difference-in-differences (DiD) approach and its limits: it is a generalization of DiD whenever important information is available as large cross-sections. After some preliminary requirements, the details of a recent piece of work in the literature are described to provide an appropriate context and notation for describing the paper's proposal.

II.1 Preliminaries

The difference-in-differences (DiD) approach usually uses ordinary least squares (OLS) in repeated cross-sections of some measure-y data of grouped individual units which are either treated or non-treated for several periods. For the sake of clarity, let's assume a complete set of T cross-sections is available (instead of just a subsequence of them) for each group $g \in \Xi \equiv \{1, 2, ..., G\}$. Every group g's temporal sequence of cross-sections is then indexed by $t \in \Upsilon \equiv \{1, 2, ..., T\}$. Let N_g be the number of individual units in each group g's sequence, so individual units are indexed by $i \in \Psi^g \equiv \{1, 2, ..., N_g\}$.

Two key assumptions are needed:

(1) The treatment is homogeneous, i.e., it is exactly the same treatment simultaneously applied to all the treated individual units, groups and time periods.

[2] The range of 4% - 7% can be taken as containing the long-run growth rate. For the Peruvian economy, the probability mass of being inside such a range has highly varied over time.

(2) The homogeneous treatment takes place instantaneously at the beginning of many periods of time $\tau \in \Gamma \subset \Upsilon$, which are the 'intervention dates'.

These assumptions allow, for the whole sample across $\{i, g, t\}$, to label many periods as 'before' $(\tau - 1)$, 'after' (τ) and even 'another period after' $(\tau + 1)$ with respect to any specific 'intervention date'.[3] This setup leads to the equation that is usually estimated to obtain the DiD estimator:

$$y_{igt} = \beta_1 \text{treat}_{igt} + \beta_2 \text{post}_{igt} + \beta_3(\text{treat}_{igt}\text{post}_{igt}) + \beta_4 X_{igt} + \alpha_0 + \alpha_g + \omega_t + \epsilon_{igt} \quad (1)$$

where $\text{treat}_{igt} = 1$ corresponds to treated individual units ($\text{treat}_{igt} = 0$, to non-treated individual units); $\text{post}_{igt} = 1$ corresponds to periods 'after' treatment ($\text{post}_{igt} = 0$, to periods 'before' treatment); X_{igt} are explanatory variables not related to the homogeneous treatment; α_0 is the intercept; α_g is the group g's fixed effect; ω_t is either the period t's fixed effect (if the available number of periods is small) or the product of a linear trend coefficient and t (if that is not the case); and ϵ_{igt} is the error term.

By defining

$$\tilde{y}_{igt} \equiv y_{igt} - (\beta_4 X_{igt} + \alpha_0 + \alpha_g + \omega_t) \quad (2)$$

equation (1) can be re-written

$$\tilde{y}_{igt} = \beta_1 \text{treat}_{igt} + \beta_2 \text{post}_{igt} + \beta_3(\text{treat}_{igt}\text{post}_{igt}) + \epsilon_{igt} \quad (3)$$

Thus, provided that $E[\epsilon_{igt}|\text{treat}_{igt}, \text{post}_{igt}] = 0$, the following expectations are obtained

$$\begin{array}{rcl} E[\tilde{y}_{igt}|\text{treat}_{igt} = 1, \text{post}_{igt} = 1] & = & \beta_1 + \beta_2 + \beta_3 \\ E[\tilde{y}_{igt}|\text{treat}_{igt} = 1, \text{post}_{igt} = 0] & = & \beta_1 \\ E[\tilde{y}_{igt}|\text{treat}_{igt} = 0, \text{post}_{igt} = 1] & = & \beta_2 \\ E[\tilde{y}_{igt}|\text{treat}_{igt} = 0, \text{post}_{igt} = 0] & = & 0 \end{array} \quad (4)$$

and by arranging them in the archetypical 2x2 matrix

Table 1: DiD estimator

	Pre (B)	Post (A)	(A-B) diff.
Treatment (T)	β_1	$\beta_1 + \beta_2 + \beta_3$	$\beta_2 + \beta_3$
Control (C)	0	β_2	β_2
(T-C) diff.	β_1	$\beta_1 + \beta_3$	β_3

β_3, the causal effect, becomes the DiD's key parameter to be estimated. β_2 could be thought as the *placebo effect*. However, while a psychological effect is not negligible when investigating the effects of a drug treatment, it should be negligible whenever the 'patient' who receives the placebo (i) does not know he/she is receiving it, and (ii) does not care about what kind of drugs the 'patient next door' is receiving. Since this is actually the case in our non-experimental discipline, an economist may consider $\beta_2 + \beta_3$ as the *direct* effect (a key component of a causal effect) whenever there is no data about individuals ('patients') not receiving any 'treatment'. To see this, consider that equation (1) becomes $y_{igt} = (\beta_2 + \beta_3)\text{post}_{igt} + \beta_4 X_{igt} + \alpha_0 + \alpha_g + \omega_t + \epsilon_{igt}$ and equation (3) becomes $\tilde{y}_{igt} = (\beta_2 + \beta_3)\text{post}_{igt} + \epsilon_{igt}$: this equation foreshadows equation (10) in subsection II.2.

[3] Set Γ is not just a subset of Υ: with just one period after treatment, its definition is $\Gamma \equiv \{\tau | \tau \in \Upsilon \wedge \tau - 1 \in \Upsilon\}$. With monthly data, having two periods after treatment allows the construction of 'experimental quarters', which imply an additional restriction in Γ's definition.

In addition to estimating all the parameters in equation (1) by OLS, the researcher can also run the following OLS regression

$$y_{igt} = \bar{\beta}_4 X_{igt} + \bar{\alpha}_0 + \bar{\alpha}_g + \bar{\omega}_t + \varepsilon_{igt} \tag{5}$$

and then use the estimated coefficients to get the estimated residuals, which can be interpreted as a corrected response (free of confounders), just like $\tilde{y}_{igt}$ in equation (2). Note that although $\tilde{y}_{igt}$ estimates include the true errors in equation (1), they correspond to the full sample and thus an appropriate division is required: divide all these 'residuals' in four sets: before-the-treatment $(\tau - 1)$ residuals for treated individual units, $\tau - 1$ residuals for non-treated individual units, after-the-treatment (τ) residuals for treated individual units, and τ residuals for non-treated individual units.[4] Then, by a direct application of the Frisch-Waugh-Lovell (FWL) theorem, there are two equivalent procedures to obtain both an estimate of the treatment effect and a test for its significance:

(i) run an OLS regression of the corresponding 4-type panel on the same explanatory dummies as in equation (3), and then use the estimate of β_3 and the corresponding standard error to built the t-test.

(ii) compute the corresponding sample means (fill the table above) as well as the sample variances $E[\tilde{y}_{igt}|\text{treat}_{igt} = a, \text{post}_{igt} = b], \quad a, b \in \{0, 1\}$, and then use all these sample moments to build the t-test for the significance of β_3. However, this solution assumes all treatments are made 'simultaneously' to all treated individual units, thus it is feasible to suppose a placebo treatment was simultaneously made to the non-treated individual units.

These details provide a framework for interpreting the literature. [Bertrand *et al.* (2004)] (BDM from now onwards) is a milestone in the literature on DiD approach for underlining severely biased standard errors because of neglected serial-correlation problems. These authors propose three techniques to solve such a problem for large sample sizes, from which the most simple one consists in *ignoring* the time series component in the estimation[5] when computing the standard errors. BDM show there are two versions of this specific technique bringing correct rejection rates and relatively high power:

(a) average the data 'before' and 'after' the treatment and then run equation (1) on the resulting averaged outcome variable as a two-period panel.[6]

(b) obtain the residuals from an auxiliary regression excluding all dummy variables associated to the treatment and divide the residuals of *the treated groups only* in two sets: before-the-treatment residuals and after-the-treatment residuals. Then proceed with an OLS regression of this two-period panel on and 'after' dummy.[7]

Note version (b) is similar to the procedure (ii) above because now it is *not* feasible to suppose a placebo treatment was simultaneously made to the non-treated individual

[4] Do not forget the 'experimental quarters' in the case of monthly data: there also exist $\tau+1$ residuals for treated individual units and $\tau + 1$ residuals for non-treated individual units.

[5] As an example, not ignoring such a component would be equivalent to postulate a common AR(1) model for each group g in equation (1), which affects the estimation strategy for all the other parameters therein.

[6] BDM note this solution works well only for treatments that are 'simultaneously' applied to all the *treated groups*. If the treatment occurs at different times for some of those groups, 'before' and 'after' are not the same for all groups and a modification is needed.

[7] BDM note this solution works as well as (a) for treatments that are 'simultaneously' applied to all the *treated groups*. Moreover, it works well when the treatments occurs at different times for some of the *treated groups*.

units, thus it is not possible to use the counterfactual information provided by non-treated individual units. Besides, there is their emphasis on *treated groups*, which will be clarified next.

II.2 Single-group tests & single-unit tests

Even though DiD approach has been pervasive in the economics literature on policy evaluation, it is not quite immune to criticism when used with observational data. Wherever the experimental setup does not hold, some drastic adaptations should be made. In general, the internal validity of model in equation (1) depends on having exactly the same treatment across different treated individual units. In the case of BDM, they explicitly take groups as states and treatment/intervention as passed laws (so that individual units may be thought as firms and the measure y, as their profits). If the law is passed in some states but not in others, then all firms in the former states will be treated and all firms in the latter states will be non-treated (by default). The model in equation (1) must then be modified as

$$y_{igt} = \beta_1 \text{treat}_{gt} + \beta_2 \text{post}_{gt} + \beta_3(\text{treat}_{gt}\text{post}_{gt}) + \beta_4 X_{igt} + \alpha_0 + \alpha_g + \omega_t + \epsilon_{igt} \qquad (6)$$

where the emphasis of the treatment has changed from individual unit i to groups g: the *treated groups* must be indexed by $g' \in \Psi \subset \Xi$. The internal validity of model in equation (6) now depends on having exactly the same passed law across different treated states/countries (groups). Otherwise, the model should be written as

$$y_{igt} = \sum_{g' \in \Psi} \beta_{1g'} \text{treat}_{g't} + \beta_2 \text{post}_{gt} + \sum_{g' \in \Psi} \beta_{3g'}(\text{treat}_{g't}\text{post}_{gt}) + \beta_4 X_{igt} + \alpha_0 + \alpha_g + \omega_t + \epsilon_{igt} \qquad (7)$$

where the assumption of simultaneous treatments still holds! This possibility is surprisingly not covered by BDM, because in their setup the analysis of state-tailored laws passed inside different states (say) should also be a reference model.[8]

The case under scrutiny here is related to both the qualitative and quantitative resources used for **the diffusion of central banks' official forecasts**. Many central banks are interested on how to use these announced forecasts to benevolently affect the private sector's expectations inside their countries, in particular those under the framework of *inflation targeting* or in the path towards passing the charter law with a clear mandate enforcing such a framework. Under these circumstances, no matter how large is the sample of 'experimental quarters', the model in equation (7) is the right setup. However, it does preclude the whole DiD approach because there is no clear counterfactual for each *treated group* $g \in \Psi$.[9] This is why the researcher is better served **by** a 'specific' model for each *treated group* $g \in \Psi$,[10]

$$\begin{gathered} y^g_{it} = \beta^g_2 \text{post}^g_t + \beta^g_4 X^g_{it} + \alpha_0 + \omega_t + \epsilon^g_{it} \\ \forall g \in \Psi \end{gathered} \qquad (8)$$

[8]BDM do make their reader note their two versions (a) and (b) of their most simple technique do poorly when the number of groups is small and it is important to mention this for our case is group $g = 1$! However, it will soon be shown that BDM's simulations are built with respect to both a model and a parameter which is different from the one this paper emphasizes.

[9]For the sake of a simplified notation, g' is abandoned from here on.

[10]Treated state in BDM or treated country in [Barrera (2018)].

from which the **single-group tests** for a singleton group[11] can be obtained by defining

$$\begin{array}{c}\tilde{y}^g_{it} = y^g_{it} - (\beta^g_4 X^g_{it} + \alpha_0 + \omega_t) \\ \forall g \in \Psi\end{array} \tag{9}$$

or by running the associated OLS regression with the whole sample for cleaning the data from the confounders' effects (an alternative analogous to the one described from equation (5) on). Then, two versions of the following equation

$$\begin{array}{c}\tilde{y}^g_{it} = \beta^g_2 \mathrm{post}^g_t + \epsilon^g_{it} \\ \forall g \in \Psi\end{array} \tag{10}$$

can be run for each *treated group* g: one for comparing the τ residuals with the $\tau - 1$ residuals and one for comparing the $\tau + 1$ residuals with the $\tau - 1$ residuals.

Thus, provided that $E[\epsilon^g_{it}|\mathrm{post}^g_t] = 0$, the following expectations are obtained

$$\begin{array}{rcl}E[\tilde{y}^g_{it}|\mathrm{post}^g_t = 1] & = & \beta^g_2 \\ E[\tilde{y}^g_{it}|\mathrm{post}^g_t = 0] & = & 0\end{array} \tag{11}$$

and by arranging them in a 1x2 matrix

Table 2: Single-group estimator

	Pre (B)	Post (A)	(A-B) diff.
Treatment (T)	0	β^g_2	β^g_2

β^g_2 becomes the single-group parameter to be estimated.

There are few steps left for reaching either the procedure in [Barrera (2018)] or the proposal in this paper: first, without information specific to firm i allowing to explain y^g_{it}, its expectation for either real growth (g) or inflation (π), X^g_{it} should be replaced by aggregate information, X^g_t; by the same token, the y^g_{it} data can then be collapsed in terms of a particular moment of the cross-section indexed by i, say, the dispersion of the cross-section of firms' expectations in country g. The single-group tests become the *single-unit tests.*[12]

II.3 Proposal

The proposal here is to collapse the y^g_{it} cross-sectional data in terms of a functional response, a probability density function, which then will allow the researcher to obtain a comprehensive list of moments by means of simulations. Specifically, the proposal requires

- to use the temporal sequence of available long-cross-sections to obtain f_t, the associated sequence of kernel-based densities;
- to use **functional regressions** to explain the evolution of the densities and to control for relevant 'confounders' (e.g., a temporal trend);
- to simulate from both the observed (f_t) and estimated ($\hat{f}_t$) densities to obtain the difference in moment r at time t, $\Delta m^r_t \equiv m^r_t(f_t) - m^r_t(\hat{f}_t)$, a ; all moments m^r are available for us to select!

[11] Since α^0 and α^g are the the coefficients associated to the same column of ones, only α^0 remains.

[12] The nonlinear regressions in [Barrera (2018)] were proposed for modeling a non-zero response such as the dispersion of expectations.

- to calculate all available differences between any Δm_t^r after a policy intervention ('treatment') and its corresponding pre-treatment, $\Delta m_{t-\Delta t}^r$. Then, built the corresponding t-tests.[13]

The literature about functional regressions provides two ways of modeling functions, that is, explaining a sequence of functions (a special variable) by means of two or more sequences of scalars (variables). The proposal uses the **fully-fledged functional approach** ([Ramsay & Silverman (1997), Ramsay & Silverman (2005)]) and the reader is referred to these books. A little warning is due here: the (alternative) **longitudinal approach** is useful when modeling sequences of continuous **sections** of demand or supply (say) at the cost of not being possible to abandon the firms' dimension i.

For illustrating the proposal above, the stylized application belongs to the literature about anchoring expectations (see footnote 1). In fact, the proposal above has [Barrera (2018)]'s methodology as its ancestor. Motivated by [Filacek & Saxa (2012)], [Barrera (2018)] used few specific scalar criteria (two robust moments) of the small cross-sections of Consensus professional forecasters' expectations to gauge the **direct** effects of *Banco Central de Reserva del Peru* (*BCRP*) forecasts. The first stage of [Barrera (2018)]'s methodology was to explain these robust moments by a relevant set of explanatory variables not related to *BCRP* forecasts (a set of confounders) and then use the estimated *errors* from those non-linear (NL) regressions[14] as the outcome variables supposedly affected by *BCRP* forecasts. Its second stage considered the chronology of *BCRP* forecasts to define 'experimental quarters' made by pre-treatment months ($s = 1$), and post-treatment months of two types: on-impact months ($s = 2$) and more-than-1-month-later month ($s = 3$), so all estimated errors of type ($s = 3$) were compared with those of type ($s = 1$) to detect significant average changes of type ($\{s = 3|s = 1\}$) by means of t-tests (Ha); the analogous procedure was followed with estimated errors of type ($s = 2$) to detect significant average changes of type ($\{s = 2|s = 1\}$). Besides, from the discussion in sub-sections II.1 and II.2, a direct effect is a gross effect, while the causal effect provided by DiD approach is a net effect; in general, these two effects are different, but in a non-experimental discipline such as Economics, these two effects can be considered the same.

While gauging the **direct** effects of *BCRP* forecasts on private expectations, it is possible to consider a different setup: a survey with large cross-sections. In the case of Peru, this data is available from *EEM*. For this case, our proposal offers a complete characterization of the **direct** effects of the availability of Central Bank's forecasts, that is, in terms of a comprehensive set of moments. For them to be consistent with each other, they should be made available from the same probability density associated to each month's cross-section. This idea naturally leads to modeling the sequence of probability densities (obtained by kernel methods) by means of a functional regression, which by following the analogy with previous paragraph, should then consider a relevant set of (scalar) explanatory variables not related to Central Bank's forecasts, etc.

Thus, the sequence of *Epanechnikov*-kernel estimated densities $\{f_t(a)\}$ is considered as as sequence of data observed without measurement noise. $f_t(a)$ is the period-t density function with domain $a \in A \equiv \{\overline{a}, \underline{a}\} \subset \Re$, $\forall t \in \{1, 2, ..., T\}$ (e.g., $a \equiv \pi$). These densities are modeled as the functional responses of a set of explanatory (scalar) variables in matrix $\boldsymbol{Z}$, $f(a) = \boldsymbol{Z} * \beta(a) + \epsilon(a)$, $\forall a \in A$:

(i) the forecasting horizon,

(ii) the level & variability of the observed variable a,

[13] One can consider two cases: a-month-after intervention effect and an on-impact intervention effect (i.e., just on time to be considered 'post-treatment'). Therefore, some special care must be taken in terms of the chronology of events. See Appendix A.

[14] One NL regression for each expectational variable (g and π) or even for each family of forecasting horizons in the available data (short-term and medium-term horizons, say).

(iii) (a lag of) the level & variability of the nominal exchange rate (FX),

(iv) the robust mean & robust standard deviation of the Consensus professional forecasters' (insiders') forecasts, and

(v) a time trend.

Explanatory variable i) is mandatory: all π & g forecasts are **fixed-event forecasts** as they refer to the end of either the current year or the next year (specific dates only). Also note application of simplified DiD approach requires *not* including the *BCRP* forecasts. For the sake of comparability, Appendix C reports the results in [Barrera (2018)] for two scalar output variables obtained from EEM cross-sections, the robust dispersions S_n and Q_n.

However, the key problem is to escape from the obviously mistaken analogy of using the estimated errors from those *functional regressions* for obtaining interpretable treatment effects. The FWL theorem can be invoked for least-squares estimation procedures of *functional regressions*.[15] Its strict application will lead to treatment effects expressed in terms of functional regression coefficients, with reduced interpretability. The solution is to use the close relationship between a general probability density and all the set of moments that can be obtained from sampling from such a density: the needed estimated errors become the differences (deltas) between the simulated moments from the observed probability densities and the simulated moments from the estimated probability densities.[16] Appendix B provides detailed information about the comprehensive list of moments used in the paper.

III Stylized Application

III.1 Data

To fully characterize the effects of Central Reserve Bank of Peru (*BCRP*)'s forecasts on Peruvian firms' expectations for real growth (g) and inflation (π), three different sources of forecasts are considered in the paper. Firstly, *BCRP* gauges private firms' expectations with a survey, the Macroeconomic Expectations Survey (*Encuesta de Expectativas Macroeconómicas* or EEM). It consists of an increasing sample of Peruvian firms who provide their forecasts for $\{g, \pi, ...\}$ to the *BCRP*'s Department of Production Activity on a monthly basis (EEM surveys' closing date is the end of the month). The EEM cross-sections of forecasts are large enough for the corresponding sequence of densities $\{f_t(a)\}$ to be non-parametrically estimated with the *Epanechnikov* kernel and immediately taken as observed data. Each element of this sequence, $f_t(a)$, is the density function of period t with domain $a \in A \equiv \{\overline{a}, \underline{a}\} \subset \Re$, $\forall t \in \{1, 2, ..., T\}$ (e.g., $a \equiv \pi$).

Secondly, *BCRP* forecasts for both variables are available from the *BCRP*'s *Inflation Reports* (IR), whose disclosure (publication and media diffusion) is made every three or four months. The IR publication defines the treatment (dichotomous) variable (the same for either π or g, one at a time) because IR publication dates define the 'experimental quarters' behind the quasi-experimental testing of the treatment effects (exact dates correspond to the press releases; see Appendix A). Single-unit t-tests for the treatment effects

[15] See [Davidson & MacKinnon (1993)]. FWL theorem can only approximately hold for other estimation procedures (e.g., generalized least squares).

[16] By simulation, there usually exists a *functional relationship* between any moment and the probability density function from which it comes. By formulae, we require the existence of a probability density function, its moment-generating function and even the moments. Then, a simple example of such a relationship would be the (robust) mean: it would be the (weighted) integral of such a probability density function. This simple idea usually holds for any (existing) moment, so the FWL theorem holds for both the *functional regression* and those simulated moments.

of *BCRP* forecasts on EEM probability density functions (*BCRP* →EEM) only use the observations inside 'experimental quarters', which are build after such an assignment of dates: 'experimental quarters' must begin with the month previous to the *IR* publication month (press release). Given that EEM surveys' closing dates follow Consensus surveys', assignment of dates for *BCRP* →EEM single-unit t-tests is almost the same (differing only for a triad of months: August 2003, March 2010 and April 2014).[17]

Finally, other explanatory variables are the robust location (median) and robust dispersions (S_n and Q_n) calculated from *Consensus Forecasts*' small cross-sections of professional forecasters' expectations about π and g in Peru.[18] Consensus Economics, Inc. asks a small sample of professional forecasters or 'insiders' (as they will be called from now on) to provide forecasts for π and g on a monthly basis. Since the closing dates of *Consensus Forecasts*' surveys is every month's 3^{rd} Monday, Appendix A defines the due precedence of *BCRP* forecasts with respect to *Consensus Forecasts*' explanatory variables (robust location and dispersions). Since EEM surveys' closing date is the end of the month, the due precedence of *BCRP* forecasts with respect to EEM *Epanechnikov* probability densities is also assured.

Besides the data and its chronology, four additional data issues need to be controlled for. Firstly, all forecasts under study are fixed-event forecasts because all of them consider two fixed events (with fixed dates): either the end of the current calendar year or the end of next calendar year. Since the maximum forecasting horizon is $H = 24$ months, the full sample of forecasts can be split into two separate sub-samples: the short-term forecasts ($h \leq 12$) and the medium-term forecasts ($12 < h \leq 24$).

The common sample of forecasts is January 2004 - December 2015. Given their fixed-event nature, this sample can only include the forecasts for the end of 2004 which were generated during the year 2004 (medium-term forecasts for the end of 2004 generated during the year 2003 are 'not available'). Similarly, this sample can only include the forecasts for the end of 2015 which were generated during the year 2015 (medium-term forecasts for the end of 2016 generated during the year 2015 are 'not available').

Secondly, there exists an important number of 'not available' data for each EEM individual firm along the monthly sample: firms can abandon the survey and then may re-enter the survey. Then, all cross-section computations (for either the EEM *Epanechnikov* densities or the EEM sample moments) only consider the available numbers, provided that EEM cross-sections are large (a similar pattern occurs for the individual insiders who provide forecasts to Consensus Economics, Inc.). The list of firms surveyed at least once has been growing fast: in January 2004, it included 432 firms, which were kept without change by January 2006; in January 2009, the list included 917 firms; in January 2012, the list included 959 firms; in March 2012, it reached 1003 firms; finally, in December 2015, the list included 1278 firms. The number of firms' plausible answers used to estimate the Epanechnikov densities has then been increasing, belonging to an approximated range of [300 500], though.

Thirdly, the EEM data first received was pre-depurated and well organized, but barely covered the last two years (2014-2015). Since the study was supposed to go back as far as January 2002, the author had to deal with non-depurated data beginning in January 2004 and ending in December 2015. The advantages of such a trade are obvious: the outlier depuration was made conservatively and homogeneously, leading to the ranges [-10 15] and [-2 15] for short-term and medium-term π expectations, respectively, and [-3 15] and

[17] Note that it is always possible to use a continuous monthly series of *BCRP* forecasts (one for π and another for g) by defining the *BCRP* forecasts as 'outstanding' (the most recently published *BCRP* forecast). This simple information-set-based strategy transforms a quarterly series into a monthly series and, in the case of the literature on mixed sampling frequencies, it provides a model which becomes a simple alternative to the Kalman filter model with missing observations in the low-frequency series (see [Foroni (2012)] and references therein).

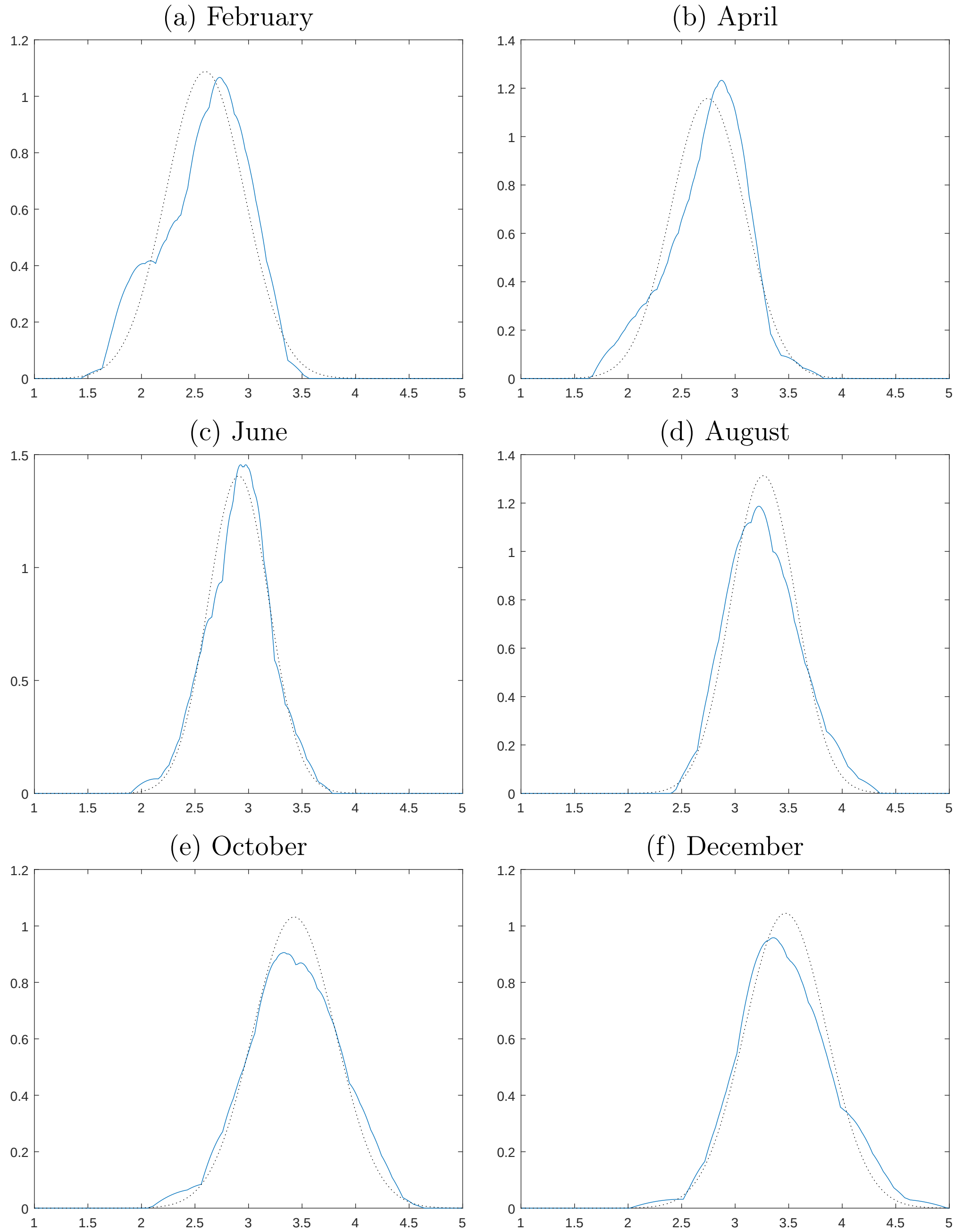
[18] See Appendix B for the definitions of the moments used in the paper.

[-1 15] for short-term and medium-term g expectations, respectively. In spite of this conservative and homogeneous data depuration, the *Epanechnikov* densities still have fat tails, so robust location (median) and dispersions (S_n and Q_n) must be considered since their means and variances may become not-well-defined in the population.

Figure 2 shows a sub-sequence of EEM densities, this time obtained from Peruvian financial entities' and analysts' short-term π (pre-depurated) expectations. This sub-sequence corresponds to an upsurge of the nominal exchange rate (FX) in Peru (beginning in August 2014). Clearly, π expectations react to nominal depreciation: the probability mass moves towards ranges of higher inflation expectations.[19] This kind of evolution clearly justifies the inclusion of (lagged) FX variables into the set of explanatory variables for the EEM

[19] The range of these economists' expectations is narrower than the ranges of the firms' expectations, which may be related to the pre-depurations made.

Figure 2: Peruvian economists' short-term π expectations during 2015



Note: Bi-monthly sequence of Epanechnikov kernel densities (continuous line) and Gaussian densities (dotted line). The graphs illustrate a recent episode of nominal FX upsurge. Besides, Epanechnikov densities are not close to their Gaussian peers (the latter densities used the sample mean and standard deviation of the same data used for obtaining the former densities).

densities of Peruvian firms' π forecasts: the monthly average and the ln(1000(standard deviation) of end-of-period daily FX interbank quotations. Lagged FX variables are needed to avoid some conceptual problems related to having two proxies of central bank credibility, one on each side of any relationship. Particular moments of these EEM densities (for instance, robust dispersions) are actually proxies of central bank credibility with respect to price stability (see [Bordo & Siklos (2015)]), and so are those FX variables.

III.2 Results

The results from short-term-horizon g expectations show that the publication of short-term g forecasts generates *on-impact* increases in the skewness, the two robust measures of kurtosis, and the probability that these expectations are in the long-run-growth-rate range of [4% 7%]. All these *on-impact* increases are consistent with *on-impact* decreases in the percentile 95, a measure of the left tail's probability mass and a measure of the right tail's probability mass. However, the *on-impact* change in the probability that these expectations are in the range has a sign opposite to the the *one-month-later* change in this probability. This *one-month-later* change is consistent with a *one-month-later* decrease in the mode and the *one-month-later* increases in the kurtosis, the robust measure of skewness and a measure of the left tail's probability mass. See Table D1 in Appendix D.

The results from short-term-horizon π expectations show that the publication of short-term π forecasts generates *on-impact* increases in robust and non-robust measures of location (trimmed means and median), as well as in the two robust measures of dispersion, the percentiles 5, 10, 15, 20, 80 & 85, and the probability that these expectations are in the target range of [1% 3%]. All these *on-impact* increases are consistent with *on-impact* decreases in the two measures of the left tail's probability mass and the percentile 95. However, some of the *one-month-later* changes have signs opposite to those *on-impact* changes (e.g., trimmed means, median, percentile 85). All these *one-month-later* decreases are consistent with a *one-month-later* increase in the skewness and a *one-month-later* decrease in the mean. See Table D2 in Appendix D.

The results from medium-term-horizon g expectations show that the publication of medium-term g forecasts generates *on-impact* increases in the two robust measures of dispersion, the percentiles 90 & 95, and the robust measure of skewness. All these *on-impact* increases are consistent with *on-impact* decreases in the percentile 15 and the probability that these expectations are in the long-run-growth-rate range of [4% 7%]. However, the *on-impact* change in the probability that these expectations are in the target range has a sign opposite to the the *one-month-later* change in this probability. This *one-month-later* change is consistent with a *one-month-later* decrease in the robust measure of skewness and a *one-month-later* increase in the non-robust measure of kurtosis. See Table E1 in Appendix E.

The results from medium-term-horizon π expectations show that the publication of medium-term π forecasts generates *on-impact* increases in robust and non-robust measures of location (mean, trimmed means, median and mode), as well as in the two robust measures of dispersion, the percentiles 5, 10, 15, 80 & 85, the robust measure of skewness and the two robust measures of kurtosis. All these *on-impact* increases are consistent with *on-impact* decreases in the two measures of the left tail's probability mass coupled with *on-impact* increases in a measure of the right tail's probability mass. However, some of the *one-month-later* changes have signs opposite to those *on-impact* changes (e.g., some location measures, percentiles 5 & 80, and the robust measure of skewness). All these *one-month-later* decreases are consistent with *one-month-later* increases in one of the robust measures of kurtosis as well as in the measure of the right tail's probability mass. Surprisingly, there are no significant changes in the probability that these expectations are in the target range of [1% 3%]. See Table E2 in Appendix E.

All these results contrast with the non-significant results from updated single-moment NL-regression-based t-tests (the robust measures of dispersion, Q_n & S_n). See Tables C1 and C2 in Appendix C.

IV Conclusions

The experimental setup and its requirements impose severe restrictions to applications where the researcher wants not only to discover whether a particular treated group g becomes significantly affected by some kind of treatment but also to explore the treated group g' conditions under which such a treatment maximizes its benevolent impact, as well as to determine specific ways to manage the treatment in the most effective way. For this kind of questions, the conditions associated to the other treated groups can bias the treatment effect because in reality there does not exist an homogeneous treatment (including their specific conditions) across treated groups (countries in our desired application).

From these problems, we build on BDM's (implicit) solution of disregarding any counterfactual. The paper provides an extension to such a solution, which allows a complete and consistent characterization of the direct effects from treatment (on-impact changes & one-month-later changes). The stylized application takes advantage from the availability of large cross-sections in EEM surveys to Peruvian firms. Benevolent effects from Peruvian Central Bank's forecasts are found for EEM firms' π expectations.

The perspectives from the empirical side are related to considering (i) the *Ha* single-unit t-tests for the short-term sample, as well as to the hypothesis of useful effects coming from Consensus forecasts, (ii) the complementary convergence data considered in [Barrera (2018)], that is, the gap between the EEM expectations and the previous *BCRP* forecasts as a new probability density function to be affected by the current *BCRP* forecasts, and (iii) the non-linear functional regressions, which will be useful for addressing relevant questions about the different direct effects of *BCRP* forecasts being above (below) the maximum (minimum) inflation allowed by the target range or just inside this range.

The perspectives from the methodological side are related to the possibility of a well-defined homogeneous and simultaneous treatment that would lead to a control set of densities (a counterfactual). In this case, a *fully-fledged DiD approach* will be feasible and our proposal will provide *full characterization* of *causal* effects of a treatment (if and only if the specific application does not allow to consider a direct effect as being the same as a causal effect). Such availability of data in terms of densities for many countries (say) would be named *huge data* instead of just *big data*.

V References

[Barrera (2018)] Barrera, Carlos (2018), Expectations and Central Banks' Forecasts: The Experience of Chile, Colombia, Mexico, Peru and the United Kingdom, 2004 - 2014, *Finance a úvěr, Czech Journal of Economics and Finance*, Vol. 68(6), 578-599.

[Bertrand *et al.* (2004)] Bertrand, Marianne; Esther Duflo & Sendhil Mullainathan (2004), How Much Should We Trust Differences-in-Differences Estimates?, *Quarterly Journal of Economics*, Vol. 119(1), 249-275.

[Blinder *et al.* (2008)] Blinder, Alan; Michael Ehrmann; Marcel Fratzscher; Jakob De Haan & David-Jan Jansen (2008), Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence, *Journal of Economic Literature*, Vol. 46(4), 910-945.

[Bonato (2011)] Bonato, Matteo (2011), Robust estimation of skewness and kurtosis in distributions with infinite higher moments, *Finance Research Letters*, Vol. 8(2), 77-87.

[Bordo & Siklos (2015)] Bordo, Michael & Pierre Siklos (2015), *Central Bank Credibility Before and After the Crisis*, NBER Working Paper No. 21710 (November).

[Brys *et al.* (2006)] Brys, Guy; Mia Hubert & Anja Struyf (2006), Robust Measures of Tail Weight, *Journal of Computational Statistics & Data Analysis*, Vol. 50(3), 733-759 (February).

[Consensus (2015)] Consensus Economics Inc. (2015), Answer to a question formulated to their editors' email address, `editors@consensuseconomics.com` (July 10, 2015).

[Croux & Rousseeuw (1992)] Croux, Christophe & Peter Rousseeuw (1992), Time-efficient Algorithms for Two Highly Robust Estimators of Scale, in Yadolah Dodge & Joe Whittaker, eds., *Computational Statistics, Volume I. Proceedings of the* 10^{th} *Symposium on Computational Statistics*, Heidelberg, Physika-Verlag, pp. 411-428.

[Davidson & MacKinnon (1993)] Davidson, Russell & James MacKinnon (1993) *Estimation and Inference in Econometrics*, Oxford University Press.

[Dräger *et al.* (2016)] Dräger, Lena; Michael Lamla & Damjan Pfajfar (2016), Are Survey Expectations Theory-Consistent? The Role of Central Bank Communication and News, *European Economic Review*, Vol. 85(June), 84-111.

[Filacek & Saxa (2012)] Filacek, Jan & Saxa, Branislav (2012) Central Bank Forecasts as a Coordination Device: Evidence from the Czech Republic, *Czech Economic Review*, Vol. 6, 244-264.

[Foroni (2012)] Foroni, Claudia (2012) *Econometric Models for Mixed-Frequency Data*, Department of Economics, European University Institute.

[Gürkaynak *et al.* (2010)] Gürkaynak, Refet; Eric Swanson & Andrew Levin (2010), Does inflation targeting anchor long-run inflation expectations? Evidence from the U.S., UK and Sweden, *Journal of the European Economic Association*, Vol. 8(6), 1208-1242.

[Hattori *et al.* (2016)] Hattori, Masazumi; Steven Kong; Frank Packer & Toshitaka Sekine (2016), *The Effects of a Central Bank's Inflation Forecasts on Private Sector Forecasts: Recent Evidence from Japan*, Bank of Japan Working Paper No. 16-E-11 (August).

[Huang *et al.* (2016)] Huang, Lei; Fabian Scheipl; Jeff Goldsmith; Jonathan Gellar; Jaroslaw Harezlak, Mathew McLean; Bruce Swihart; Luo Xiao; Ciprian Crainiceanu & Philip Reiss (2016), *R Package* **`Refund`***: Regression with Functional Data*, Version 0.1-14.

[Kozicki & Tinsley (2005)] Kozicki, Sharon & P.A. Tinsley (2005), What do you expect? Imperfect Policy Credibility and Test of the Expectations Hypothesis, *Journal of Monetary Economics*, Vol. 52(2), 421-447 (March).

[Kumar *et al.* (2015)] Kumar, Saten; Hassan Afrouzi; Olivier Coibion; & Yuriy Gorodnichenko (2015), *Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand*, NBER Working Paper 21814 (December).

[Neuenkirch (2013)] Neuenkirch, Matthias (2013), Central Bank Transparency and Financial Market Expectations: The Case of Emerging Markets, *Economic Systems*, Vol. 37(), 598-609.

[Pereira da Silva (2016)] Pereira da Silva, Luis Awazu (2016), *Old and New Challenges for 2016 and Beyond: Strengthening Confidence by Re-anchoring Long-term Expectations*, paper presented as the Lamfalussy Lecture Series, at the Professor Lamfalussy Commemorative Conference 'His Contribution to Economic Policy and the Birth of the Euro', Budapest (February).

[Pedersen (2015)] Pedersen, Michael (2015), What Affects the Predictions of Private Forecasts? The Role of Central Bank Forecasts in Chile, *International Journal of Forecasting*, Vol. 31(4), 1043-1055 (October-December).

[Ramsay & Silverman (1997)] Ramsay, James & R.W. Silverman (1997), *Functional Data Analysis*, Springer.

[Ramsay & Silverman (2005)] Ramsay, James & R.W. Silverman (2005), *Functional Data Analysis*, Second Edition, Springer.

[Rousseeuw & Croux (1993)] Rousseeuw, Peter & Christophe Croux (1993), Alternatives to the Median Absolute Deviation, *Journal of the American Statistical Association*, Vol. 88(424), 1273-1283 (December).

[Sigman (2010)] Sigman, Karl (2010), *Inverse Transform Method*, Class notes, [Retrieved 2019-11-11] Available at: `http://www.columbia.edu/~ks20/4404-Sigman/4404-Notes-ITM.pdf`.

[Trabelsi (2016)] Trabelsi, Emna (2016), Central Bank Transparency and the Consensus Forecast: What Does The Economist Poll of Forecasters Tell Us?, in *Research in International Business and Finance*, Vol. 38(June), 338-359.

Appendices

A Chronologies

Table A1: Assignment of *BCRP* forecasts to Consensus Economics Inc.'s surveys ∗/

Return to Section III.1 or Subsection II.3

Dates associated with Peru's *IR*s					
Number	*IR*	Press Release	*IR* tentative assignment from *LACF* survey 1/	*LACF Survey Date* close to the Press Release 2/	*IR* final assignment from *LACF* survey 1/
	Aug03	29aug03	(Sep03)	18aug03	(Sep03)
1	Jan04	06feb04	Feb04	16feb04	Feb04
2	May04	04jun04	Jun04	21jun04	Jun04
3	Aug04	10sep04	Sep04	20sep04	Sep04
4	Jan05	04feb05	Feb05	21feb05	Feb05
5	May05	03jun05	Jun05	20jun05	Jun05
6	Aug05	02sep05	Sep05	19sep05	Sep05
7	Jan06	03feb06	Feb06	20feb06	Feb06
8	May06	02jun06	Jun06	19jun06	Jun06
9	Sep06	06oct06	Oct06	16oct06	Oct06
10	Jan07	09feb07	Feb07	19feb07	Feb07
11	May07	08jun07	Jun07	18jun07	Jun07
12	Sep07	05oct07	Oct07	15oct07	Oct07
13	Jan08	08feb08	Feb08	18feb08	Feb08
14	May08	13jun08	Jun08	16jun08	Jun08
15	Sep08	10oct08	Oct08	20oct08	Oct08
16	Mar09	13mar09	Mar09	16mar09	Mar09
17	Jun09	12jun09	Jun09	15jun09	Jun09
18	Sep09	18sep09	Oct09	21sep09	Sep09
19	Dec09	18dec09	Jan10	14dec09	Jan10
20	Mar10	26mar10	Apr10	15mar10	Apr10
21	Jun10	18jun10	Jul10	21jun10	Jun10
22	Sep10	17sep10	Oct10	20sep10	Sep10
23	Dec10	17dec10	Jan11	13dec10	Jan11
24	Mar11	18mar11	Apr11	21mar11	Mar11
25	Jun11	17jun11	Jul11	20jun11	Jun11
26	Sep11	16sep11	Oct11	19sep11	Sep11
27	Dec11	16dec11	Jan12	19dec11	Dec11
28	Mar12	23mar12	Apr12	19mar12	Apr12
29	Jun12	15jun12	Jun12	18jun12	Jun12
30	Sep12	14sep12	Sep12	17sep12	Sep12
31	Dec12	14dec12	Dec12	17dec12	Dec12
32	Mar13	22mar13	Apr13	18mar13	Apr13
33	Jun13	21jun13	Jul13	17jun13	Jul13
34	Sep13	20sep13	Oct13	16sep13	Oct13
35	Dec13	20dec13	Jan14	16dec13	Jan14
36	Apr14	25apr14	May14	22apr14	May14
37	Jul14	18jul14	Aug14	21jul14	Jul14
38	Oct14	17oct14	Nov14	20oct14	Oct14
39	Jan15	23jan15	Feb15	19jan15	Feb15
40	May15	22may15	Jun15	18may15	Jun15
41	Sep15	18sep15	Oct15	14sep15	Oct15
42	Dec15	18dec15	Jan16	14dec15	Jan16

∗/ Consensus survey's closing date is always before EEM's (the end of the month).

1/ *Consensus Economics Inc.* carries out the Latin-American-country survey every month's 3^{th} Monday ([Consensus (2015)]). A tentative assignment of the central bank *IR* forecasts to the *Consensus Economics Inc.* surveys considers that these forecasts will surely affect the survey's forecasts from the very month of an *IR* publication (until they become affected by the following *IR*'s forecasts) if the *IR* publication date falls before or at the 14th day of that month; otherwise, they will surely affect the survey from the following month to the publication month (until they become affected by the following *IR*'s). The final assignment uses the closing date of the corresponding *Consensus Economics Inc.*'s survey.

2/ For the case of the effects upon the *EEM*'s forecasts, both *Consensus Economics Inc.*'s dates and *IR* Press Releases' dates indicate that these two types of forecasts will contemporaneously affect the *EEM*'s forecasts (except maybe for March 2010's *IR*). While the frequency of *Consensus Economics Inc.*'s forecasts is monthly (allowing a direct use of the auxiliary regression), Central Bank's *IR* forecasts still require a specially-tailored 'assignment' similar to the one used in the previous paper.

B Moments

The simulations are obtained from each estimated density corresponding to period $t \in \{1, 2, ..., T\}$, thus allowing to obtain a comprehensive set of scalar moments for each estimated density:

1. First-order moments: mean; 5%- and 10%-trimmed means;[20] median (percentile $\{50\}$); and mode.

2. Second-order moments: standard deviation; robust dispersion estimators Q_n and S_n proposed by [Rousseeuw & Croux (1993)].[21]

3. Higher-order moments: skewness, SK_2; kurtosis, KR_2, KR_4;

4. Other moments: $Pr(range)$,[22] its confidence interval and its variability coefficient; Percentiles $\{5, 10, 15, 20, 80, 85, 90, 95\}$; LQW_s and LQW_b left tails ($s = 0.125$ and $b = 0.250$); RQW_s and RQW_b right tails ($s = 0.875$ and $b = 0.750$).

Some clarifications are due regarding some 'other moments'. Traditional standardized moments such as skewness (SK)[23] and kurtosis (KR)[24] actually depend upon other traditional moments like the mean or the variance, which may not exist in the population's distribution. Sample counterparts are always computable, but their values will then display an erratic behavior; see [Bonato (2011)]. Corresponding robust measures SK_2, KR_2 and KR_4 are preferred,

$$SK_2 \equiv \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$KR_2 \equiv \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2}$$

$$KR_4 \equiv \frac{F^{-1}(0.975) - F^{-1}(0.025)}{F^{-1}(0.750) - F^{-1}(0.250)}$$

where Q_i is the i-th quartile,[25] and E_i is the i-th octile, that is, $E_i \equiv F^{-1}(i/8)$ for $i \in \{1, 2, ..., 7\}$; see [Bonato (2011)]. Before continuing with the specificities of the simulations, note KR, KR_2 and KR_4 have two statistical disadvantages: (i) they are really measuring not only the tail heaviness but also the peakedness of a distribution, and (ii) their tail-heaviness interpretation is restricted to symmetric distributions. [Brys *et al.* (2006)] recommend the use of robust measures of left and right tails, the left quantile weight (LQW_p), and the right quantile weight RQW_q (for $0 < p < \frac{1}{2}$ and $\frac{1}{2} < q < 1$, respectively).

$$LQW_p \equiv \frac{F^{-1}(\frac{1-p}{2}) + F^{-1}(\frac{p}{2}) - 2F^{-1}(0.250)}{F^{-1}(\frac{1-p}{2}) - F^{-1}(\frac{p}{2})}$$

[20] The p% trimmed mean of n sampled values $\{x_1, x_2, ..., x_n\}$ is the mean of those values excluding the highest and lowest q data values, where $q = n * (p/100)/2$.

[21] Given a sample of n points, $\{x_1, x_2, ..., x_n\}$, $S_n \equiv s_{mp} s_{mg} med_i\{med_j\{|x_i - x_j|\}\}$ and $Q_n \equiv q_{mp} q_{mg}\{|x_i - x_j|; \quad i < j\}_{(k)}, \quad k \equiv \binom{h}{2}, \quad h \equiv \lfloor n/2 \rfloor + 1$, where $\{y_i\}_{(k)}$ refers to the k-th order statistic obtained from the data set $\{y_i\}$; $\binom{a}{b}$, to the combinations of a elements taken in groups of b elements; and $\lfloor c \rfloor \equiv \max\{d \in \mathbb{Z} | d \leq c\}$, to the maximum integer of c. s_{mg} and q_{mg} are the adjustment factors compensating for the (asymptotic) large-sample bias with respect to a normal distribution, and s_{mp} and q_{mp}, the adjustment factors compensating for the small-sample bias; see [Croux & Rousseeuw (1992)].

[22] The scalar criterion $Pr(range)$ is the probability that the variable defining the support of the densities (functional responses in the functional regression model) happens to be inside the 'range'. This range is [4 7] for g forecasts and [1 3] for π forecasts.

[23] If SK is positive [negative], the long tail is to the right [left].

[24] If KR is regarded as a measure of tail heaviness, a positive [negative] KR means a symmetric distribution has heavier tails [lighter tails] than a normal distribution's tails.

[25] Given a process of n points, $\{x_1, x_2, ..., x_n\}$, and assuming that the x_j's are independent and identically distributed with cumulative distribution function F, $Q_1 \equiv F^{-1}(0.25)$, $Q_2 \equiv F^{-1}(0.50)$, and $Q_3 \equiv F^{-1}(0.75)$

$$RQW_p \equiv \frac{F^{-1}(\frac{1+q}{2}) + F^{-1}(1-\frac{q}{2}) - 2F^{-1}(0.750)}{F^{-1}(\frac{1+q}{2}) - F^{-1}(1-\frac{q}{2})}$$

C $\boldsymbol{Ha}$ t-tests for BCRP $\rightarrow \boldsymbol{EEM}$

Table C1: Tests with Q_n dispersion of EEM forecasts

Return to Section III.2

Variable	**Model/d.f.**	Current vs. Previous ($\{s=2\|s=1\}$)		Next vs. Previous ($\{s=3\|s=1\}$)	
		$Tcal$	p_1 ($p-value$)	$Tcal$	p_2 ($p-value$)
		Short-term sample ($h \leq 12$)			
GDP growth	Add.trend/32	-0.275	0.393	-0.262	0.398
CPI inflation	Add.trend/32	0.226	0.411	-0.692	0.247
		Medium-term sample ($h > 12$)			
GDP growth	Add.trend/31	0.186	0.427	-0.052	0.479
CPI inflation	Add.trend/31	0.049	0.480	-0.023	0.491

See [Barrera (2018)]'s Online Appendix, Table E.1.

Table C2: Tests with S_n dispersion of EEM forecasts

Return to Section III.2

Variable	**Model/d.f.**	Current vs. Previous ($\{s=2\|s=1\}$)		Next vs. Previous ($\{s=3\|s=1\}$)	
		$Tcal$	p_1 ($p-value$)	$Tcal$	p_2 ($p-value$)
		Short-term sample ($h \leq 12$)			
GDP growth	Add.trend/32	-1.193	0.121	-0.515	0.305
CPI inflation	Add.trend/32	-0.372	0.356	-0.792	0.217
		Medium-term sample ($h > 12$)			
GDP growth	Add.trend/31	0.254	0.401	-0.045	0.482
CPI inflation	Add.trend/31	-0.260	0.398	-0.147	0.442

See [Barrera (2018)]'s Online Appendix, Table D.1.

D *Ha* t-tests for BCRP →*EEM* (moment-simulated deltas)

Table D1: BCRP →EEM (g)

Return to Section III.2

Ha t-tests for EEM-moment-simulated deltas (m1, short-term sample, $h \leq 12$)

Variable	**Simulated Scalar Criteria** (moments)	Current vs. Previous ($\{s=2\|s=1\}$) $Tcal$ 34 d.f.	p_1 ($p-value$)	Next vs. Previous ($\{s=3\|s=1\}$) $Tcal$ 34 d.f.	p_2 ($p-value$)
GDP growth	Mean	1.000	0.162	-1.000	0.162
	Trimmean5	-0.737	0.233	-0.520	0.303
	Prctile50	-0.431	0.335	-0.455	0.326
	Mode *	-0.406	0.344	-1.762	**0.044**
	Std.Dev.	-1.000	0.162	1.000	0.162
	Skewness	1.771	**0.043**	-0.381	0.353
	Kurtosis	-0.960	0.172	1.728	**0.047**
	Prctile5	0.928	0.180	-1.148	0.130
	Prctile10	0.563	0.289	-1.231	0.114
	Prctile15	-0.157	0.438	-1.267	0.107
	Prctile20	-0.430	0.335	-0.929	0.180
	Prctile80	0.019	0.492	0.840	0.204
	Prctile85	-0.645	0.262	0.855	0.199
	Prctile90	-1.165	0.126	0.971	0.169
	Prctile95	-1.347	**0.094**	0.994	0.164
	Trimmean10	-0.556	0.291	-0.659	0.257
	SK_2	0.740	0.232	1.319	**0.098**
	KR_2	1.909	**0.033**	1.300	0.101
	KR_4	1.379	**0.089**	0.820	0.209
	LQW_s	0.906	0.186	1.527	**0.068**
	LQW_b	-1.475	**0.075**	0.686	0.249
	RQW_s	0.656	0.258	-0.512	0.306
	RQW_b	-1.774	**0.043**	0.983	0.166
	Q_n	0.819	0.209	0.629	0.267
	S_n	0.951	0.174	0.595	0.278
	$ub\{Pr(.)\}$ ‡	5.149	**0.000**	-1.769	**0.043**
	Pr(range) †	5.134	**0.000**	-1.753	**0.045**
	$lb\{Pr(.)\}$ ‡	5.119	**0.000**	-1.737	**0.046**
	$cv\{Pr(.)\}$	-4.180	**0.000**	-1.781	**0.042**

* Not simulated. †g & π ranges: [4 7] & [1 3]. ‡Pr(range)'s 95% CI.

Table D2: BCRP →EEM (π)

Return to Section III.2

Ha t-tests for EEM-moment-simulated deltas (m1, short-term sample, $h \leq 12$)

Variable	**Simulated Scalar Criteria** (moments)	Current vs. Previous ($\{s=2\|s=1\}$)		Next vs. Previous ($\{s=3\|s=1\}$)	
		$Tcal$ 32 d.f.	p_1 ($p-value$)	$Tcal$ 31 d.f.	p_2 ($p-value$)
CPI inflation	Mean	1.067	0.147	-1.552	**0.065**
	Trimmean5	3.516	**0.001**	-2.436	**0.010**
	Prctile50	1.342	**0.095**	-1.815	**0.044**
	Mode *	-0.037	0.485	-1.179	0.124
	Std.Dev.	0.000	0.500	1.129	0.134
	Skewness	-1.253	0.110	1.661	**0.053**
	Kurtosis	0.976	0.168	-0.690	0.248
	Prctile5	2.338	**0.013**	-0.219	0.414
	Prctile10	3.608	**0.001**	-0.609	0.273
	Prctile15	2.097	**0.022**	-0.877	0.194
	Prctile20	1.699	**0.050**	-1.081	0.144
	Prctile80	4.085	**0.000**	-1.046	0.152
	Prctile85	2.858	**0.004**	-1.510	**0.071**
	Prctile90	0.566	0.288	-1.104	0.139
	Prctile95	-1.866	**0.036**	-1.127	0.134
	Trimmean10	3.832	**0.000**	-2.226	**0.017**
	SK_2	1.225	0.115	-0.394	0.348
	KR_2	-0.417	0.340	-0.065	0.474
	KR_4	-0.989	0.165	0.267	0.396
	LQW_s	-1.386	**0.088**	0.197	0.422
	LQW_b	-3.897	**0.000**	0.788	0.218
	RQW_s	0.381	0.353	-0.270	0.394
	RQW_b	1.006	0.161	-1.307	0.100
	Q_n	1.603	**0.059**	0.099	0.461
	S_n	2.672	**0.006**	0.174	0.431
	$ub\{Pr(.)\}$ ‡	2.405	**0.011**	0.818	0.210
	Pr(range) †	2.419	**0.011**	0.819	0.210
	$lb\{Pr(.)\}$ ‡	2.433	**0.010**	0.820	0.209
	$cv\{Pr(.)\}$	-0.600	0.276	-0.021	0.492

* Not simulated. †g & π ranges: [4 7] & [1 3]. ‡Pr(range)'s 95% CI.

E *Ha* t-tests for BCRP →*EEM* (moment-simulated deltas)

Table E1: BCRP →EEM (g)

Return to Section III.2

Ha t-tests for EEM-moment-simulated deltas (m2, medium-term sample, $h > 12$)

Variable	**Simulated Scalar Criteria** (moments)	Current vs. Previous ($\{s=2\|s=1\}$) $Tcal$ 37 d.f.	p_1 ($p-value$)	Next vs. Previous ($\{s=3\|s=1\}$) $Tcal$ 37 d.f.	p_2 ($p-value$)
GDP growth	Mean	0.583	0.282	0.753	0.228
	Trimmean5	0.381	0.353	0.274	0.393
	Prctile50	0.168	0.434	-0.276	0.392
	Mode *	0.427	0.336	-0.562	0.289
	Std.Dev.	-0.207	0.418	1.116	0.136
	Skewness	0.640	0.263	-0.632	0.266
	Kurtosis	-0.891	0.189	1.925	**0.031**
	Prctile5	-0.064	0.475	0.935	0.178
	Prctile10	-1.061	0.148	0.445	0.329
	Prctile15	-1.353	**0.092**	0.085	0.466
	Prctile20	-1.179	0.123	-0.295	0.385
	Prctile80	1.074	0.145	-0.241	0.405
	Prctile85	1.236	0.112	-0.154	0.439
	Prctile90	1.622	**0.057**	-0.775	0.222
	Prctile95	2.316	**0.013**	-1.153	0.128
	Trimmean10	0.241	0.406	0.055	0.478
	SK_2	-1.452	**0.077**	-1.664	**0.052**
	KR_2	0.048	0.481	-0.592	0.279
	KR_4	-0.948	0.175	0.135	0.447
	LQW_s	-1.010	0.160	0.982	0.166
	LQW_b	-0.950	0.174	-0.665	0.255
	RQW_s	1.153	0.128	-1.029	0.155
	RQW_b	1.119	0.135	0.495	0.312
	Q_n	2.164	**0.019**	0.717	0.239
	S_n	2.079	**0.022**	0.902	0.187
	$ub\{Pr(.)\}$ ‡	-1.320	**0.097**	1.568	**0.063**
	Pr(range) †	-1.326	**0.096**	1.576	**0.062**
	$lb\{Pr(.)\}$ ‡	-1.333	**0.095**	1.583	**0.061**
	$cv\{Pr(.)\}$	1.694	**0.049**	-1.752	**0.044**

* Not simulated. †g & π ranges: [4 7] & [1 3]. ‡Pr(range)'s 95% CI.

Table E2: BCRP →EEM (π)

Return to Section III.2

Ha t-tests for EEM-moment-simulated deltas (m2, medium-term sample, $h > 12$)

Variable	**Simulated Scalar Criteria** (moments)	Current vs. Previous ($\{s=2\|s=1\}$) $Tcal$ 35 d.f.	p_1 ($p-value$)	Next vs. Previous ($\{s=3\|s=1\}$) $Tcal$ 34 d.f.	p_2 ($p-value$)
CPI inflation	Mean	3.907	**0.000**	0.959	0.172
	Trimmean5	3.750	**0.000**	-1.760	**0.044**
	Prctile50	2.268	**0.015**	-1.326	**0.097**
	Mode *	1.262	0.108	0.105	0.458
	Std.Dev.	0.346	0.366	-0.960	0.172
	Skewness	1.119	0.135	0.268	0.395
	Kurtosis	-0.114	0.455	-0.867	0.196
	Prctile5	1.783	**0.042**	-2.684	**0.006**
	Prctile10	1.906	**0.032**	-0.966	0.171
	Prctile15	1.686	**0.050**	-1.146	0.130
	Prctile20	1.115	0.136	-0.957	0.173
	Prctile80	4.276	**0.000**	-1.475	**0.075**
	Prctile85	2.595	**0.007**	-1.289	0.103
	Prctile90	-0.278	0.391	-1.434	**0.080**
	Prctile95	-0.869	0.195	-1.715	**0.048**
	Trimmean10	3.996	**0.000**	-1.571	**0.063**
	SK_2	1.738	**0.046**	-1.894	**0.033**
	KR_2	1.918	**0.032**	2.313	**0.013**
	KR_4	2.214	**0.017**	-0.114	0.455
	LQW_s	-2.344	**0.012**	0.098	0.461
	LQW_b	-1.706	**0.048**	-0.415	0.340
	RQW_s	1.000	0.162	0.376	0.355
	RQW_b	1.747	**0.045**	2.033	**0.025**
	Q_n	2.289	**0.014**	-0.850	0.201
	S_n	2.774	**0.004**	-0.868	0.196
	$ub\{Pr(.)\}$ ‡	-1.012	0.159	1.090	0.142
	Pr(range) †	-1.016	0.158	1.084	0.143
	$lb\{Pr(.)\}$ ‡	-1.020	0.157	1.079	0.144
	$cv\{Pr(.)\}$	-0.003	0.499	-0.283	0.390

* Not simulated. †g & π ranges: [4 7] & [1 3]. ‡Pr(range)'s 95% CI.