(Non)rationality of consumer inflation perceptions

Peter, Eckley

University of Oxford

29 January 2015

(Non)rationality of consumer inflation perceptions

Peter Eckley
Department of Economics
University of Oxford

Supervised by Professor Chris Bowdler

Abstract:

We test the rationality of consumer inflation perceptions in Sweden, relaxing assumptions that have been maintained in previous literature. Specifically, we test the rational expectations hypothesis on survey measures of inflation perceptions, interpreted as nowcasts. We progressively relax restrictions on the prior set of loss functions against which the perception errors may be rationalised, culminating in the first application to inflation perceptions of the indicator test of Patton & Timmerman (2007). We find that inflation perceptions are positively biased in both the short- and long-run and thus reject RPH. This contrasts with earlier literature using data from the same survey. We also find that inflation perceptions fail to efficiently incorporate even the information implicit in past perception errors.

JEL codes: D12, D84, E31

Keywords: inflation, perceptions, expectations, rational expectations hypothesis, forecast rationality, nowcasting

Contact: peter.eckley@gmail.com

Table of Contents

1 Introduction

Since the rational expectations revolution of the 1970's the assumption that economic agents form their expectations rationally – as if solving the economist's model, and making efficient use of all available information – has become a cornerstone of the dominant economic modelling paradigm. An enormous literature has empirically tested the so-called rational expectations hypothesis (REH). Forward-looking inflation expectations have been particularly widely studied in light of inflation's role as a nominal anchor in modern monetary policy frameworks. The literature has generally rejected at least strong versions of REH.

Perceptions on the other hand have attracted relatively little empirical work. However, a nascent empirical literature has highlighted that consumer inflation perceptions differ substantially, significantly and systematically from inflation outturns (e.g. Curto Millet, 2006; Jonung & Laidler, 1988; 2011). This has serious but rarely acknowledged implications for both macroeconomic models and the quantification of forward-looking inflation expectations from qualitative survey responses, both of which typically assume more benign properties for inflation perception errors.

This paper tests the rationality of consumer inflation perceptions in Sweden over the 15-year span from 1993m1 to 2007m12, using the population-weighted mean of natively quantitative survey responses. Inflation perceptions can be conceptualised as 'nowcasts' of inflation because official figures are published with a lag of one to two months relative to the survey. We thus formalise the rational perceptions hypothesis (RPH) as the special case of REH with zero forecast horizon. We then test observable implications of RPH for inflation perception errors, steadily relaxing restrictions on the prior set of loss functions against which the perception errors may be rationalised, culminating in the first application to inflation perceptions of the indicator test of Patton & Timmerman (2007) which allows us to relax restrictions that have been maintained in previous literature.

We find that Swedish inflation perceptions are positively biased in both the short- and long-run and thus reject RPH. This contrasts with some earlier literature using data from the same survey (Jonung & Laidler, 1988; Lein & Maag, 2011). We also find that inflation perceptions fail to efficiently incorporate even the information implicit in past perception errors.

The remainder of this thesis is structured as follows. Section 2 formally states RPH in a loss-theoretic framework. Section 3 introduces the data. Section 4 tests RPH under a range of assumptions on the loss function and DGP. Section 5 draws conclusions and compares our results to the extant literature. Section 6 highlights some important implications of our findings for recent theory models of inflation perceptions, macroeconomic modelling, and the quantification of forward-looking inflation expectations that are important in monetary policy.

2 The rational perceptions hypotheses

Our starting point is the recognition that perceptions may be viewed as a special case of expectations, where the forecast horizon is zero periods into the future. Therefore tests of the rational expectations hypothesis (REH) can be applied to perceptions.

Since Muth's (1961) original formulation[1], REH has since been reframed in terms of loss functions and formalised as follows. For the variable to be forecasted at time t over horizon h, x_{t+h}, there exists some loss function, L^*, among the prior set of loss functions $\mathcal{L}$ that are considered possible/plausible *a priori*, relative to which the forecast, $\tilde{x}_{t+h}$, minimises expected loss conditional on a given information set, Ω_t. More concisely:

$$\text{REH: } \exists\, L^* \in \mathcal{L} \text{ s.t. } \tilde{x}_{t+h} = argmin_{\check{x}_{t+h}}\, E_t[L^*(x_{t+h}, \check{x}_{t+h})|\Omega_t]$$

where $E_t[\cdot\,|\Omega_t]$ denotes the expectations operator at time t conditional on Ω_t. It is typically assumed that the forecaster knows the conditional distribution of $x_{t+h}|\Omega_t$ so that their expectations operator coincides with that of the stochastic process generating x_{t+h}. The rational perceptions hypothesis (RPH) is then simply the special case with forecast horizon h equal to zero.

In this study, the variable to be forecasted (or 'nowcasted') is $x_t = \pi_t^n \equiv P_t/P_{t-n} - 1$, the actual n–month inflation rate in the reference index P up to month t, and $\tilde{x}_t = \tilde{\pi}_t^n$ is the perception formed, in month t, of π_t^n. The available perceptions data is reported monthly and corresponds to annual inflation rates, i.e. $n = 12$, but the more general n notation will aid discussion of the econometric methodology below.

RPH is always a joint hypothesis on the three entities underlined above – reference index, information set, conditional distribution, and the prior set of loss functions – in addition to the behavioural assumption of minimising expected loss and the assumption of knowledge of the unconditional distribution. Assumptions on these entities are stated and justified in conjunction with the testing methodology in Section 4.

3 The data

3.1 Why Sweden?

Our monthly Swedish data are particularly appropriate for the purposes of our study for four reasons.

First, to the best of our knowledge, this is the only natively *quantitative* (as opposed to directional/qualitative) survey data on inflation perceptions at a national scale that is publicly available. This circumvents the need to impose questionable distributional assumptions (which become maintained hypotheses throughout testing) to arrive at quantified perception figures that can be cardinally compared to actual inflation.

Second, the 15-year span covered by this series is longer than similar series at sub-national scale or unpublished series[2].

Third, the survey question (see below) mentions "inflation" explicitly, as well as "prices in general" making it more likely that respondents use an headline inflation index as their reference index rather than, for example, the prices of salient individual goods such as petrol or milk. Furthermore, CPI has been the headline official inflation index in Sweden since 1954 so the potential for confusion from changing official indices (as in the UK from RPI to CPI for example) does not arise.

---

[1] "The subjective probability distribution of outcomes tends to be distributed, for the same information set about... 'objective' probability distributions of outcomes [predicted by the economic theory]".

[2] The Inflation Psychology Survey started in August 1997 and is restricted to 500 households in Ohio, USA (M. F. Bryan & Venkatu, 2001a, 2001b). The European Commission initiated a quantitative survey in 2003 but had not published the data at the time of writing.

Fourth, monetary policy in Sweden was remarkably stable throughout the sample period. Just before 1993m1 when our sample period begins Sweden transitioned from a fixed- to floating-exchange rate regime and announced the introduction of inflation targeting (to begin in 1995m1), to which it adhered beyond the end of the sample (2007m12). There have been no radical shifts in the exchange rate during the period (see Table 9 in the Appendix for a more detailed monetary policy timeline) in contrast to countries that joined the Euro and experienced substantial impacts on inflation perceptions following cash changeover in 2002 (Ehrmann, 2006). As a result, Bryan and Palmqvist (2005) find that inflation during this period was well approximated as lying in a single regime.

The main shortcoming of our data, which is perhaps the flipside of the stability just mentioned, is that inflation was relatively low and stable by historical and international standards during our sample period (in the range −1.2 to +5.2 percentage points), so the data can shed little light on the behaviour of inflation perceptions when inflation is very high, as it was in the 1970s for example.

3.2 Sample period

The start of the sample was dictated by the first publication of monthly perceptions data. Data for an earlier quarterly survey was not available to us.

The end of the sample was dictated by the most recent available data published at the time we undertook the original analysis. During further work on this paper data for later years became available. We chose not to extend the sample because our exploratory analysis found evidence of a structural break in early 2008. Dias, Duarte, & Rua (2010) similarly finds an inflation regime shift in 2008 in Eurozone countries, and assesses it to be even more important than the break arising from the cash changeover to Euro in 2002 in Eurozone countries. The authors thus end their sample at 2007m12, as do Lein & Maag (2011) and Dräger (2014).

3.3 Variable definitions

Inflation is calculated from the non-seasonally adjusted monthly shadow CPI index numbers[3]. Inflation perceptions ($\tilde{\pi}_t^{12}$) data derive from responses to a question in the monthly Consumer Tendency Survey[4]:

"Compared with 12 months ago, how much higher in percent do you think that prices are now – in other words, the present rate of inflation?"

The only distributional statistics published by SCB are the mean of all responses, and the mean excluding extreme responses which are defined as <-5% or >15% (Konjunktur Institutet, n.d.). and which we estimate may account for around 5% of responses[5].

[3] 'Shadow CPI' is the most up-to-date revised figure for that same month at the time we downloaded the data (December 2008). By comparing perceptions to shadow CPI we are implicitly assuming that rational consumers attempt to nowcast 'true' inflation (of which shadow CPI is the best available estimate) rather than the first official estimate of CPI figure *per se*. In practice, this makes negligible difference to our analysis because CPI revisions are occasional and small (as can be seen in Figure 1), usually correcting for minor administrative errors. See SCB (2001) for further details.

[4] Formerly called the Consumer Survey, and before that the Households' Purchasing Plans survey.

[5] According to Figure 3 of Palmqvist & Strömberg (2004) less than 4% of respondents reported inflation *expectations* above 15%, and less than 1% below -5% (although data is only reported for 2001m11 and 2001m12). It seems plausible that similar proportions would hold for *perceptions*, given that mean inflation perceptions closely track expectations (see Appendix, Figure 2).

Both means are reported after re-weighting to be representative of the Swedish adult population (aged 16 to 84), with the exclusion of extreme responses done before re-weighting. The population mean is a natural empirical counterpart to the representative household found in most macroeconomic models. That said neither mean is likely to be an ideal location statistic for the empirical cross-sectional distribution of responses because inflation perceptions are commonly found to be non-normally distributed[6]. We therefore report test results for both measures as a robustness check.

Figure 1 shows π_t^{12}, $\tilde{\pi}_t^{12}$ and perception errors $\pi_t^{12} - \tilde{\pi}_t^{12}$ over time.

3.4 Order of integration

The empirical evidence on the order of integration of inflation is mixed in the literature (e.g. Altissimo, Ehrmann, & Smets, 2006) and in our sample (see Appendix). On subsamples of our Swedish data the null of a unit root in inflation perceptions has not been rejected in recent literature (e.g. Dräger, 2014; Lein & Maag, 2011). However, non-rejection does not imply acceptance of the null and the fact that inflation consistently remains in a relatively narrow range over 15 years suggests stationarity. We proceed in a manner that accommodates either conclusion.

3.5 Relative timing of survey and information publication

The price data underlying CPI was collected during the week that the 15th of the month. The corresponding CPI was published between 7th and 22nd of the following month[7]. The Consumer Tendency Survey was conducted over a multi-week period. For example, during 2003 (the only year for which precise dates were available) data gathering covered the first 20 days of each month. Thus the most recently published CPI figures available to a respondent surveyed in month t could be for either $t-2$ or $t-1$.

3.6 Intervention dummies

We use two intervention dummies:

$$\mathrm{d93} = \begin{cases} 1 \text{ if } t = 1993\mathrm{m}1, \dots, 1993\mathrm{m}12 \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

$$\mathrm{dgfk} = \begin{cases} 1 \text{ if } t \geq 2002\mathrm{m}1 \\ 0 \text{ if } t < 2002\mathrm{m}1 \end{cases} \tag{2}$$

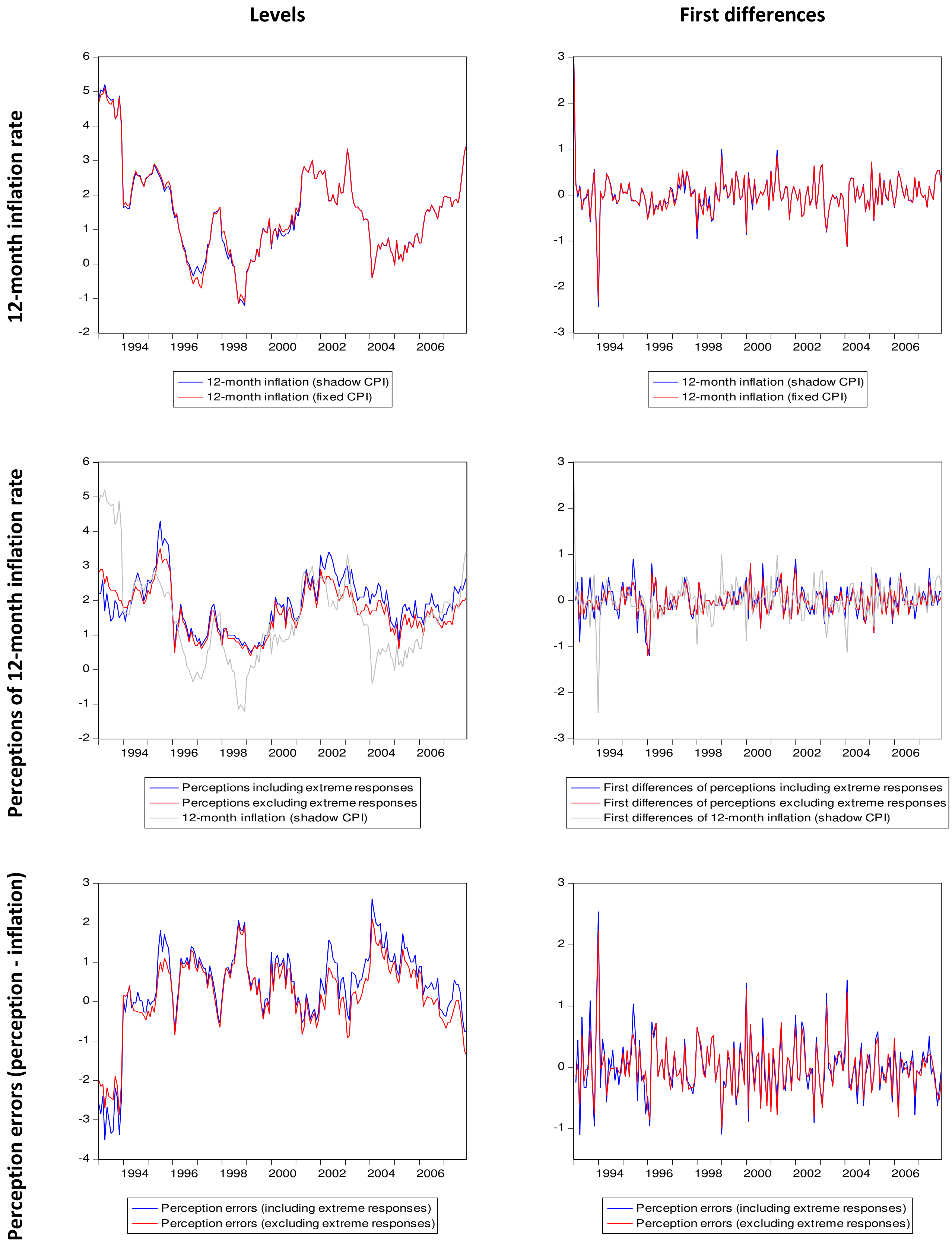
d93 covers the first year following the announcement of inflation targeting when inflation perceptions were relatively flat around the target level of 2% (though this was not due to come into effect officially until 1995m1) while a large spike in vehicle fuel prices drove CPI inflation up to around 5%, resulting in repeated large negative perception errors. Whether this reflects backward-looking perceptions based on 1992 (when inflation was around 1.5 to 2%) or anchoring to the policy target or some other cause, d93 prevents these outliers dominating the model fit, while retaining information from the intra-year variation.

dgfk attempts to control for a change, in 2002m1, of the organisation conducting the survey, from SCB to GfK, which resulted in estimates of $\tilde{\pi}_t^{12}$ around one percentage point higher during 2001m11-m12 when both surveys were run concurrently. Palmqvist and Strömberg (2004) found such a dummy to be less than entirely satisfactory, so we also test on subsamples 1993m1–2001m12 and 2002m1–2007m12.

---

[6] For example Bruine de Bruin et al. (2012) finds substantial right skew and excess kurtosis in the cross-sectional distribution of consumer inflation perceptions in other European countries.
[7] Source: http://www.scb.se/Pages/Standard____33842.aspx, downloaded on 2 April 2009

Figure 1: Inflation and inflation perceptions, levels and first differences, Sweden, 1993m1–2007m12



4 Testing RPH

In this Section we report several tests of RPH, interleaving econometric methodology and results.

Most of the results in this section concern properties of perception errors, which are I(0) under the null of RPH regardless of order of integration of π_t^{12}, so that test statistics have the standard stationary distributions under the null. The exception is the Mincer-Zarnowitz regression, from which we want to estimate the perceptions bias under the alternative hypothesis. This is discussed in Section 4.2.

To deduce testable implications of RPH we must first specify our assumptions about the reference index, information set, conditional distribution and prior set of loss functions, which are jointly part of RPH as highlighted in Section 2. We maintain fixed assumptions about the reference index and information set throughout, as outlined in Section 4.1, and we focus on testing RPH under and increasingly general prior set of loss functions and conditional distributions.

4.1 Maintained assumptions

The reference index is assumed throughout to be the official CPI inflation rate introduced in Section 3. We focus on testing RPH under and increasingly general prior set of loss functions and conditional distributions.

Where an assumption on the information set is necessary for testing (which is not always the case) then we assume that it contains (but is not necessarily restricted to) π_{t-s}^{10} , i.e. the part of $\pi_t^{12} \approx \pi_t^2 + \pi_{t-2}^{10}$ that is publicly available throughout the period during which the survey for month t is conducted (see Section 3.5). Therefore under RPH the problem is essentially to 'nowcast' π_t^2 based on the information set. We do not attempt to assess rationality with respect to the 'ragged edge' of incomplete information on current and immediate past values of inflation that may be publicly observable ahead of the announcement of official CPI as discussed in Wallis (1986), nor do we concern ourselves with private information.

The assumptions regarding information set and conditional distribution are in general intimately related. The 'conditional' in conditional distribution of π_t^{12} refers to conditioning on the information set. Changing the information set – in particular adding or removing information relevant to nowcasting inflation – could change the conditional distribution. Thus, in general, a rejection of RPH obtained under a particular restrictions on the information set and conditional distribution, only implies rejection of RPH under that pair of restrictions.

Nevertheless for all combinations of assumptions on the loss function and conditional distribution that are tested in this paper, a rejection of RPH using a particular information set Ω also implies rejection of RPH for any information set that is a superset of Ω, so long as the assumptions (if any) on the corresponding conditional distribution still hold. This is because all the combinations tested below imply exclusion restrictions (in a regression context) or orthogonality restrictions (with respect to perception errors) for all items of the information set, under the null of RPH. For example, under MSE or MAE loss, RPH implies unbiased perceptions without making any assumptions on Ω or the conditional distribution (see below), so that if we can reject unbiasedness then this rejection holds for empty Ω and thus for any information set, which could include the 'ragged edge' and arbitrary private information, and for any conditional distribution.

4.2 Mean-square error (MSE) loss

If the loss function is of MSE form (i.e. $L \propto (\tilde{\pi}_t^{12} - \pi_t^{12})^2$) then RPH implies i) perceptions are mean-unbiased (i.e. expected perception error is zero $E_t[\tilde{\pi}_t^{12} - \pi_t^{12}] = 0$) and ii) perception errors are uncorrelated with any element of the information set. This follows without any restrictions on the conditional distribution or information set. Clearly this is very restrictive assumption on the loss function, but one widely used in economic modelling.

Since the information set contains π_{t-2}^{10}, any perception errors in $\tilde{\pi}_t^{12}$ must relate to the π_{t-1}^{1} and π_t^1 components of $\pi_t^{12} \approx \pi_t^1 + \pi_{t-1}^1 + \pi_{t-2}^{10}$. The π_{t-1}^1 component is also contained in π_{t-1}^{12} and subject to perception error in $\tilde{\pi}_{t-1}^{12}$ so that perception errors may be correlated at the first lag. However, neither π_{t-1}^1 nor π_t^1 are contained in π_{t-s}^{12} $\forall s \geq 2$, so that under RPH perception errors should be uncorrelated at the second and longer lags.

Mean-unbiasedness can be tested parametrically via a Wald test of $H_0: \delta = 0$ versus $H_a: \delta \neq 0$ in the regression

$$\tilde{\pi}_t^{12} - \pi_t^{12} = \delta + \eta_t \tag{3}$$

or via a Wald test of $H_0: \alpha = 0$ & $\beta = 1$ versus $H_a: Not\ H_0$ in the traditional rationality regression of Mincer and Zarnowitz (Mincer & Zarnowitz, 1969)

$$\pi_t^{12} = \alpha + \beta\tilde{\pi}_t^{12} + \varepsilon_t \tag{4}$$

where the error terms η_t and ε_t are MA(1) under H_0 because of the potential serial correlation of perception errors at the first lag but not at longer lags, as discussed above.

The test in (3) is more precise in that $\alpha = 0$ & $\beta = 1$ in (4) constitutes a sufficient but not necessary condition for unbiasedness, which is satisfied more generally by $\alpha = (1 - \beta)E[\pi_t^{12}]$ (Holden & Peel, 1990).

We estimate (4) in addition because it has become standard in the literature and is interpretable as a linear approximation to the conditional expectation function of π_t^{12} conditional on $\tilde{\pi}_t^{12}$. Note that the parameter restrictions on (4) mentioned above only test RPH if π_t^{12} and $\tilde{\pi}_t^{12}$ and I(0). If instead they are I(1) and cointegrated (tested by a DF-GLS test of the null of a unit root in the residuals) then (4) recovers the cointegrating vector asymptotically, so we are effectively testing equality between π_t^{12} and $\tilde{\pi}_t^{12}$ in the long-run equilibrium, but the dynamics of perception errors as they tend towards that equilibrium are unconstrained (apart from the restriction of stationarity). However, if RPH holds then it holds in the long-run equilibrium as well, so that a rejection of $H_0: \alpha = 0$ & $\beta = 1$ would imply rejection of RPH, even though a non-rejection could be consistent with RPH.

We estimate (3) and (4) using OLS (consistent under H_0) both with and without the dummies d93 and dgfk introduced in Section 3.5 and label the corresponding parameters θ_{93} and θ_{gfk} where the dummies are included. Recursive parameter estimates in our parametric regressions (not reported for the sake of space) showed no serious instability problems except early in the sample if the d93 dummy was excluded.

Standard errors should allow for heteroskedasticity, since the null of RPH does not imply homoskedastic errors. However, regarding residual serial correlation a subtlety arises. Under the null the residuals are MA(1). Therefore to test RPH as a package of assumptions our standard error estimates need only allow for MA(1) structure. However, to test unbiasedness separately from the other implications of RPH (such as no residual serial correlation) we must allow for unrestricted serial correlation in our standard error estimates. (Indeed the residuals exhibit strong autocorrelation at lags longer than 1, and partial autocorrelations are better approximated by AR(1) than MA(1).) In both cases we HAC-robust estimator of Newey and West (1987)[8], but in the MA(1) (null) case we truncate the kernel after the first lag to obtain more accurate standard error estimates and more power to reject the null than in the second case where in the face of *a priori* unknown serial correlation set the truncation lag to the function of sample size suggested by Newey and West (1994)[9].

Zero serial correlation perception errors beyond the first lag can be tested as zero serial correlation beyond the first lag in the residuals from (3) and (4) [10], using the method of Bartlett (1946) to estimate significance levels on autocorrelation and partial autocorrelation coefficients.

Finally, a Wald test of $\beta = 1$ in (4) is interpreted as a test of a necessary condition for information efficiency, against the alternative $\beta \neq 1$ which implies information *in*efficiency in the sense that π_t^{12} could then be used to reduce the variance of the perception error $\mathrm{V}[\tilde{\pi}_t^{12} - \pi_t^{12}] = \mathrm{V}[\eta_t]$[11]. More intuitively, given the interpretation of (4) as a linear approximation to the conditional expectation function of π_t^{12} conditional on $\tilde{\pi}_t^{12}$, $\beta = 1$ implies that on average $\tilde{\pi}_t^n$ tracks the *movement* in π_t^n, even if the level is subject to a non-zero offset.

---

[8] The covariance matrix estimator of Hansen and Hodrick (1980), originally proposed to deal with overlapping data cases like ours, imposes homoscedasticity and is not guaranteed to be positive definite.

[9] Four lags in samples 1993m1–2007m12 and 1993m1–2001m12, and three lags for 2002m1–2007m1.

[10] Lein & Maag (2011) and Dräger (2014) consider serial correlation only at the 12th lag. This avoids the challenges of overlapping errors, but sacrifices efficiency and power.

[11] To see this, subtract π_t^{12} from both sides of (4), take the variance of both sides, and observe that with $\beta \neq 1$ the residual variance in the regression $\mathrm{V}[\varepsilon_t]$ is smaller than $\mathrm{V}[\eta_t]$ (Clements & Hendry, 1998).

Table 1: Bias of inflation perceptions, estimated by OLS on equation (3)

Survey measure →		Including extreme responses						Excluding extreme responses					
Sample →		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations,T →		180		108		72		180		108		72	
No dummies													
Constant	$\hat{\delta}$	0.377		0.151		0.716		0.175		0.075		0.326	
	i)	(0.112)***	[0.001]	(0.164)	[0.360]	(0.110)***	[0.000]	(0.094)*	[0.064]	(0.137)	[0.585]	(0.113)***	[0.005]
	ii)	_(0.166)**_	_[0.024]_	_(0.245)_	_[0.539]_	_(0.145)***_	_[0.000]_	_(0.139)_	_[0.208]_	_(0.202)_	_[0.712]_	_(0.148)**_	_[0.031]_
Residual normality		***	[0.000]	***	[0.000]		[0.720]	***	[0.000]	***	[0.000]		[0.905]
Residual ARCH		***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]
Residual	Lag	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC
autocorrelation	1	**0.880**	**0.880**	**0.888**	**0.888**	**0.805**	**0.805**	**0.874**	**0.874**	**0.883**	**0.883**	**0.835**	**0.835**
and partial	2	**0.775**	0.005	**0.793**	0.020	**0.627**	-0.061	**0.751**	-0.056	**0.774**	-0.030	**0.657**	-0.134
autocorrelation	3	**0.702**	0.084	**0.718**	0.047	**0.536**	0.142	**0.673**	0.123	**0.690**	0.057	**0.586**	**0.255**
With dummies													
Constant	$\hat{\delta}$	0.526		0.526				0.370		0.370			
	i)	(0.090)***	[0.000]	(0.090)***	[0.000]			(0.087)***	[0.000]	(0.087)***	[0.000]		
	ii)	_(0.125)***_	_[0.000]_	_(0.125)***_	_[0.000]_			_(0.123)***_	_[0.003]_	_(0.123)***_	_[0.003]_		
d93	$\hat{\theta}_{93}$	-3.372		-3.372				-2.659		-2.659			
		(0.141)***	[0.000]	(0.141)***	[0.000]			(0.113)***	[0.000]	(0.113)***	[0.000]		
Dgfk	$\hat{\theta}_{gfk}$	0.190						-0.045					
		(0.143)	[0.184]			*As above*		(0.143)	[0.754]			*As above*	
Residual normality			[0.276]		[0.110]				[0.796]		[0.541]		
Residual ARCH		***	[0.000]	***	[0.000]			***	[0.000]	***	[0.000]		
Residual	Lag	AC	PAC	AC	PAC			AC	PAC	AC	PAC		
autocorrelation	1	**0.761**	**0.761**	**0.724**	**0.724**			**0.802**	**0.802**	**0.769**	**0.769**		
and partial	2	**0.600**	0.049	**0.571**	0.098			**0.642**	-0.003	**0.619**	0.066		
autocorrelation	3	**0.485**	0.031	**0.426**	-0.037			**0.550**	0.101	**0.496**	0.002		

Notes: Two sets of Newey and West (1987) HAC standard errors reported for parameters to be tested: i) truncating the kernel after first lag, suitable for testing null of unbiasedness jointly with MA(1) residual structure implied by RPH; ii) selecting truncation lag based on sample size as per Newey and West (1994), suitable for testing null of unbiasedness independently of residual dependence structure (*in italics*). P-values in brackets [].* indicates significance at 10% level, ** at 5%, *** at 1%. Null of 'residual normality' tested by the omnibus test of Doornik and Hansen (2008). Null of no residual ARCH effects at lag 1 tested using Engle's (1982) LM test. Bold type for AC and PC coefficients indicates magnitude $\geq 2/\sqrt{T}$ (equal to 0.149 for T=180 observations, 0.192 for T=108, 0.236 for T=72) which is an approximate 5% significance level Bartlett (1946).

Table 2: Mincer-Zarnowitz regression tests of unbiasedness and no residual serial correlation, using OLS on equation (4) (excluding intervention dummies)

Survey measure	→	Including extreme responses						Excluding extreme responses					
Sample	→	1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T	→	180		108		72		180		108		72	
Regressor	Coef.												
Constant	$\hat{\alpha}$	-0.312		-0.317		-0.670		-0.936		-0.960		-0.619	
	i)	(0.243)	[0.200]	(0.328)	[0.335]	(0.341)*	[0.053]	(0.229)***	[0.000]	(0.273)***	[0.001]	(0.321)*	[0.057]
	ii)	*(0.341)*	*[0.361]*	*(0.467)*	*[0.499]*	*(0.420)*	*[0.116]*	*(0.324)****	*[0.005]*	*(0.389)***	*[0.015]*	*(0.403)*	*[0.129]*
Perceptions,	$\hat{\beta}$	0.966		1.097		0.978		1.453		1.541		1.169	
$\tilde{\pi}_t^{12}$	i)	(0.114)***	[0.000]	(0.185)***	[0.000]	(0.159)***	[0.000]	(0.154)***	[0.000]	(0.197)***	[0.000]	(0.178)***	[0.000]
	ii)	*(0.156)****	*[0.000]*	*(0.261)****	*[0.000]*	*(0.189)****	*[0.000]*	*(0.223)****	*[0.000]*	*(0.289)****	*[0.000]*	*(0.216)****	*[0.000]*
R^2		0.312		0.350		0.386		0.563		0.624		0.384	
Wald tests													
<u>Mean-unbiasedness</u>:		5.72***	[0.004]	0.68	[0.511]	21.04***	[0.000]	11.97***	[0.000]	7.12***	[0.001]	4.93**	[0.010]
$\hat{\alpha} = 0$ & $\hat{\beta} = 1$	F(2,T-2)	*2.61**	*[0.076]*	*0.32*	*[0.731]*	*12.07****	*[0.000]*	*6.34****	*[0.002]*	*3.82***	*[0.025]*	*2.76**	*[0.070]*
<u>Tracks movements</u>:		0.09	[0.764]	0.28	[0.601]	0.02	[0.892]	8.71***	[0.004]	7.54***	[0.007]	0.90	[0.345]
$\hat{\beta} = 1$	F(1,T-1)	*0.05*	*[0.825]*	*0.14*	*[0.711]*	*0.01*	*[0.910]*	*4.14***	*[0.043]*	*3.50**	*[0.064]*	*0.61*	*[0.438]*
Residual normality		***	[0.000]	***	[0.000]		[0.727]	***	[0.001]	**	[0.024]		[0.922]
Residual ARCH		***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]
Residual	Lag	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC
autocorrelation	1	**0.882**	**0.882**	**0.880**	**0.880**	**0.807**	**0.807**	**0.823**	**0.823**	**0.815**	**0.815**	**0.824**	**0.824**
and partial	2	**0.779**	0.002	**0.781**	0.026	**0.629**	-0.064	**0.662**	-0.048	**0.653**	-0.035	**0.640**	-0.119
autocorrelation	3	**0.706**	0.084	**0.702**	0.044	**0.538**	0.141	**0.573**	0.128	**0.537**	0.045	**0.578**	**0.271**

Notes: see notes to Table 1.

Table 3: Mincer-Zarnowitz regression tests of unbiasedness and no residual serial correlation, using OLS on equation (4) (including intervention dummies)

Survey measure	→	Including extreme responses						Excluding extreme responses					
Sample	→	1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T	→	180		108		72		180		108		72	
Regressor	Coef.												
Constant	$\hat{\alpha}$	-0.520		-0.530		-0.670		-0.640		-0.643		-0.619	
	i)	(0.190)***	[0.007]	(0.225)**	[0.020]	(0.341)*	[0.053]	(0.181)***	[0.001]	(0.208)***	[0.003]	(0.321)*	[0.057]
	ii)	_(0.260)**_	_[0.047]_	_(0.311)*_	_[0.092]_	_(0.420)_	_[0.116]_	_(0.240)**_	_[0.011]_	_(0.289)**_	_[0.028]_	_(0.403)_	_[0.129]_
Perceptions,	$\hat{\beta}$	0.997		1.002		0.978		1.176		1.177		1.169	
	i)	(0.103)***	[0.000]	(0.127)***	[0.000]	(0.159)***	[0.000]	(0.103)***	[0.000]	(0.123)***	[0.000]	(0.178)***	[0.000]
$\tilde{\pi}_t^{12}$	ii)	_(0.141)***_	_[0.000]_	_(0.261)***_	_[0.000]_	_(0.189)***_	_[0.000]_	_(0.144)***_	_[0.000]_	_(0.173)***_	_[0.000]_	_(0.216)***_	_[0.000]_
d93	$\hat{\theta}_{93}$	3.373***	(0.103)	3.372***	(0.146)			2.499***	(0.149)	2.498***	(0.206)		
			[0.000]		[0.000]				[0.000]		[0.000]		
Dgfk	$\hat{\theta}_{gfk}$	-0.189	(0.151)					0.008	(0.140)				
			[0.377]						[0.952]				
R^2		0.748		0.827		0.386		0.768		0.852		0.384	
Wald tests													
Mean-unbiasedness:		16.94***	[0.000]	16.92***	[0.000]	21.04***	[0.000]	10.99***	[0.000]	10.56***	[0.000]	4.93**	[0.010]
$\hat{\alpha} = 0$ & $\hat{\beta} = 1$	F(2,T-2)	_8.78***_	_[0.000]_	_8.78***_	_[0.000]_	_12.07***_	_[0.000]_	_6.01***_	_[0.003]_	_5.76***_	_[0.004]_	_2.76*_	_[0.070]_
Tracks movements:		0.00	[0.975]	0.00	[0.985]	0.02	[0.892]	2.90*	[0.091]	2.09	[0.151]	0.90	[0.345]
$\hat{\beta} = 1$	F(1,T-1)	_0.00_	_[0.982]_	_0.00_	_[0.989]_	_0.01_	_[0.910]_	_1.49_	_[0.224]_	_1.05_	_[0.308]_	_0.61_	_[0.438]_
Residual normality			[0.279]		[0.110]		[0.727]		[0.825]		[0.610]		[0.922]
Residual ARCH		***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]	***	[0.000]
Residual	Lag	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC	AC	PAC
autocorrelation	1	**0.762**	**0.762**	**0.723**	**0.723**	**0.807**	**0.807**	**0.769**	**0.769**	**0.719**	**0.719**	**0.824**	**0.824**
and partial	2	**0.601**	0.050	**0.570**	0.097	**0.629**	-0.064	**0.584**	-0.018	**0.530**	0.026	**0.640**	-0.119
autocorrelation	3	**0.486**	0.031	**0.424**	-0.038	**0.538**	0.141	**0.483**	0.099	**0.383**	-0.013	**0.578**	**0.271**

Notes: see notes to Table 1.

The point estimates $\hat{\delta}$ of perceptions mean-bias in Table 1 are positive and substantial (ranging between 0.33 and 0.72 percentage points, cf. an inflation target of 2%) across both perception measures. When the d93 dummy is included the bias is significant at the 1% level across all samples. Essentially the same pattern of results is seen in the mean-unbiasedness tests in the Mincer-Zarnowitz regressions in Table 2 and Table 3, which relax the restriction implicit in Table 1 that $\beta = 1$.

However, the significance of the bias is not robust to the exclusion of the d93 dummy, which is a potential concern because this was included for statistical rather than persuasive institutional reasons. It is therefore reassuring that the nonparametric tests in Section 4.3 reject median-unbiasedness at the 5% level even including 1993 and no analogue to d93. This is also consistent with mean-unbiasedness if the conditional distribution of inflation is symmetric – an assumption with which the non-rejection of normality tests is consistent. The dgfk dummy is insignificant but the point estimate is non-trivial and we retain this dummy as a precaution.

Information efficiency in the sense of no residual autocorrelation beyond the first lag is roundly rejected by the autocorrelation coefficients (AC) in Table 1, Table 2 and Table 3. These show significant autocorrelation up to at least the 4th lag (reporting truncated at the 3rd lag in the Tables for the sake of space) and in most cases up the 6th or 7th lag. Including dummies d93 and dgfk only slightly attenuates the residual autocorrelation, and combined with the relative parameter stability this makes it seem unlikely that the residual autocorrelation is spuriously driven by un-modelled structural breaks.

Our looser sense of information efficiency $\beta = 1$ is also rejected when extreme responses are excluded, but not when they are included, in which case the point estimate $\hat{\beta}$ close to 1.

The R^2 values are smaller than might be expected under RPH given that 10 or 11 of the 12 months of inflation data is already publicly available, so that perception errors should only relate to the inflation component in the final one or two months out of 12. (The apparently impressive R^2>0.8 in Table 3 is in considerable part due to the d93 dummy.)

If π_t^{12} and $\tilde{\pi}_t^{12}$ are I(1) and cointegrated then our estimates of (4) can be interpreted as a linear approximation to the long-run equilibrium, as discussed above. The DF-GLS test of Elliott, Rothenberg, & Stock (1996) **–** with lag length of one selected by information criteria, no deterministic trend, and d93 and dgfk dummies included – rejects the null of a unit root in the residuals at the 95% level for the perceptions measure including extreme responses (test statistic equal to -3.610 vs. critical value -1.950) and at the 90% level for the perceptions measure excluding extreme responses (test statistic equal to -1.857 vs. critical value -1.615). This is evidence that if π_t^{12} and $\tilde{\pi}_t^{12}$ are I(1) then they are also cointegrated. We thus conclude that if the data is I(1) then the bias estimated above persists even in the long-run.

Finally, notice the strong evidence of residual heteroskedasticity the parametric regression results. While OLS point estimates and Newey-West standard errors are asymptotically consistent in the presence of heteroskedasticity (or non-normality), they are not efficient and may not perform well in finite samples. This provides one motivation for the non-parametric tests in Section 4.3.

4.3 Mean absolute error (MAE) loss

Similar to the MSE case above, if the loss function is of MAE form (i.e. $L \propto |\tilde{\pi}_t^{12} - \pi_t^{12}|$) then RPH implies i) perceptions are *median*-unbiased (i.e. perception errors have zero mean) and ii) perception errors are uncorrelated with any element of the information set. This follows without any restrictions on the conditional distribution or information set.

As in Section 4.2, given the maintained assumption on the information set, ii) implies that that perception errors may be serially correlated at the first lag but not at longer lags.

We follow Campbell & Ghysels (1995) in applying non-parametric tests for i) and ii). These are exact in the presence of non-normality and heteroskedasticity. They are more robust to potential finite sample problems, including outliers, than regression-based tests. Parametric tests could also be conducted using quantile regression, but would be less robust.

Note that, in the special case where the conditional distribution of inflation is symmetric, then under MSE (or indeed any symmetric loss function) the optimal distribution of perception errors is symmetric (as can be easily seen by symmetry arguments) so that the mean and median coincide. In this case the test rejections and the parametric estimates of the bias in Section 4.2 can also be interpreted as applicable under MAE loss. Also the tests in this Section can be interpreted as applicable under MSE loss in which case they are more robust and potentially more efficient tests of RPH under MSE loss.

The null that perception errors have zero median and are uncorrelated beyond the first lag can be tested using statistics, S_{odd} and S_{even}, defined as

$$\begin{aligned} S_i &= \sum_{t \in T_i} u(\tilde{\pi}_t^{12} - \pi_t^{12}), i = \{odd, even\} \\ u(z) &= 1, \qquad \text{if } z \geq 0 \\ &= 0, \qquad \text{if } z < 0 \\ T_{odd} &= \{1, 3, 5, \dots\}, \qquad T_{even} = \{2,4,6, \dots\} \end{aligned} \tag{5}$$

are each binomially distributed with number of trials equal to the cardinality of T_{odd} and T_{even} and probability of success 0.5. We reject the null at significance level α if either S_{odd} or S_{even} is significant at the $\alpha/2$ level. The odd/even split ensures that, under the null, the observations involved in S_i are uncorrelated, which is assumed in calculating critical values.

Under the additional assumption that perception errors are symmetric about zero (which could arise from both loss function and DGP being symmetric, or from particular combinations of asymmetric loss function and asymmetric DGP) W_{odd} *and* W_{even}, defined as

$$\begin{aligned} W_i &= \sum_{t \in T_i} u(\tilde{\pi}_t^{12} - \pi_t^{12}) \cdot R_{1,t}^{+} \\ R_{1,t}^{+} &= \sum_{s \in T_i} u\left(|\tilde{\pi}_t^{12} - \pi_t^{12}| - |\tilde{\pi}_s^{12} - \pi_s^{12}|\right) \\ i &\in \{odd, even\} \end{aligned} \tag{6}$$

(with $u(\cdot)$, T_{odd} and T_{even} defined as before) are each distributed as Wilcoxon (1945) signed rank[12]. We reject the null at significance level α if either W_{odd} or W_{even} is exceeds the critical value for $\alpha/2$. Critical values and p-values are obtained from the normal approximation with continuity correction, as implemented in Stata's -signrank- command.

There is no straightforward analogue to the dummies used in the regression tests. In place of d93 we report results for subsamples starting in 1994m1. Instead of dgfk we examine subsamples where this is zero and one respectively.

The null of no serial correlation in perception errors at given lag l can be tested using nonparametric sign and Wilcoxon signed-rank tests due to Dufour (1981) under the maintained hypothesis that forecast errors are median-unbiased:

$$SC_l = \sum_{t=l+1}^{T} u\left((\tilde{\pi}_t^{12} - \pi_t^{12}) \cdot \left(\tilde{\pi}_{t-l}^{12} - \pi_{t-l}^{12}\right)\right) \tag{7}$$

and

$$\begin{aligned} WC_l &= \sum_{t=l+1}^{T} u\left((\tilde{\pi}_t^{12} - \pi_t^{12}) \cdot \left(\tilde{\pi}_{t-l}^{12} - \pi_{t-l}^{12}\right)\right) \cdot R_{2,t}^{+} \\ R_{2,t}^{+} &= \sum_{t=l+1}^{T} u\left(|\tilde{\pi}_t^{12} - \pi_t^{12}| \cdot |\tilde{\pi}_s^{12} - \pi_s^{12}|\right) \end{aligned} \tag{8}$$

Under the null of no residual serial correlation at lag l, SC_l is binomially distributed with number of trials equal to T and probability of success 0.5; WC_l is distributed as Wilcoxon signed-rank.

The null is strongly rejected in all cases, except for a marginal non-rejection of median-zero perception errors for the 1993m1-2001m12 subsample excluding extreme responses. We conclude that the median bias is non-zero, perception errors are serially correlated and we reject RPH for MAE loss functions.

[12] Note that for tied observations $R_{1,t}^{+}$ assigns the highest rank associated with each tie. Dufour (1981) acknowledges that using the average rank would be more natural but does not implement this for reasons of computational convenience that are no longer a material concern. We use the Stata implementation in the command -signrank- that assigns average ranks to ties. The difference is negligible in continuous data like ours where true ties occur with probability zero.

Table 4: Non-parametric sign and signed-rank tests of unbiasedness and no residual serial correlation in inflation perceptions

Survey measure →	Including extreme responses										Excluding extreme responses									
Sample →	1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1994m1–2007m12		1994m1–2001m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1994m1–2007m12		1994m1–2001m12	
Observations, T →	180		108		72		168		96		180		108		72		168		96	
Unbiasedness	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i	S_i	W_i
$i = odd$	0.000	0.000	0.040	0.020	0.000	0.000	0.000	0.000	0.002	0.000	0.015	0.004	0.220	0.103	0.029	0.010	0.001	0.000	0.029	0.000
$i = even$	0.000	0.000	0.020	0.029	0.000	0.000	0.000	0.000	0.001	0.000	0.004	0.008	0.220	0.154	0.004	0.010	0.000	0.000	0.029	0.001
Joint signif.	***	***	**	**	***	***	***	***	***	***	***	***			***	***	***	***	**	***
Residual serial correlation	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l	SC_l	WC_l
Lag, l 1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
3	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
4	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.006**	**0.000**
5	**0.000**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	0.030	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.125	**0.004**
6	**0.000**	**0.000**	**0.000**	**0.000**	**0.003**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.001**	**0.000**	0.037	**0.000**	**0.006**	**0.000**	**0.003**	**0.000**	0.125	**0.005**

Notes: S_i, W_i, SC_l, WC_l are defined in equations (5), (6), (7), (8) respectively. p-values (computed using the normal approximation with continuity correction) are reported for probability that test statistics are drawn from distribution implied by the null. * indicates acceptance of the null of median-unbiasedness at 10% level, ** at 5%, *** at 1%. p-values of SC_l and WC_l below the 1% significance threshold are highlighted in bold.

4.4 Unknown loss function

Of course empirically the loss function is unknown in our context and if it departs from MSE or MAE forms then except in special cases we require restrictions on the conditional distribution of the target variable (here inflation) to obtain testable properties of the perception errors.

By restricting the *conditional* distribution we are implicitly invoking a particular information set. However, testing can still be conducted without access to or use of that information set. Rather the interpretation of accepting or rejecting the test of properties is predicated on the restrictions on the distribution conditional on the particular information set, whatever the conditioning information set happens to be.

As noted above, if we restrict the conditional distribution to be symmetric around its mean then we can relax the restriction on the loss function to simply require symmetry and again obtain the implication of mean zero perception errors.

However, symmetric loss functions such as MSE and MAE arguably do not provide the most plausible description of many aspects of consumer behaviour (Granger & Newbold, 1986, p. 125; Zellner, 1986), particularly where there may be loss aversion, which implies a loss function that is asymmetric around zero. While it may not be clear what a 'loss' is with regard to consumer inflation perceptions[13], it is clear that imposing symmetry is restrictive and could be responsible for spurious rejections of RPH if incorrect.

Patton & Timmerman (Patton & Timmermann, 2007) (PT from hereon) show how to trade off restrictions on the loss function against restrictions on the data-generating process to obtain testable properties of perception errors under more general classes of loss functions, including many asymmetric loss functions.

If dynamics in the distribution of inflation are restricted to the mean, and the loss function is solely a function of forecast error (but subject to no other restrictions) then, as in Sections 4.2 and 4.3, RPH implies that perception errors are orthogonal to the information set and thus exhibit no residual serial correlation beyond the first lag. Hence we can extend the interpretation of the above tests of this property. Restricting inflation dynamics to the mean looks implausible against historical data in the long run, but it might reasonably approximate inflation dynamics over shorter spans.

If the restrictions on the DGP are weakened to allow dynamics in the conditional variance of inflation, and the restrictions on the loss function are tightened to require a homogeneous function, then PT shows that i) the rational perception is equal to the same quantile of the inflation distribution in every period, and ii) an indicator variable I_t for positive perception error (i.e. $I_t = 1$ if $\tilde{\pi}_t^{12} - \pi_t^{12} > 0$ and zero otherwise) is distributed independently of elements in the information set.

[13] Idiosyncratic features of individuals' loss functions might motivate asymmetric loss. For example, inflation below expectations implies a higher real-terms value of debt stock, which translates into a real-terms gain, relative to expectations, for net creditors but a real-terms loss for net debtors. If we suppose these idiosyncrasies 'average out' in the representative mean perception captured by our data, then macroeconomic non-linearities (e.g. the risk of deflationary and hyperinflationary spirals) and social norms might motivate asymmetry.

These properties are also implied by the pair of assumptions above (dynamics restriction to the mean and loss function solely a function or forecast error).

Remarkably these properties can be tested by parameter restrictions $H_0\colon \boldsymbol{\beta} = \mathbf{0}\ \forall i$ vs $H_a\colon \mathrm{Not}\ H_0$ in the simple regression:

$$g(I_t) = \alpha + \boldsymbol{\beta}'\boldsymbol{Z}_t + \eta_t \tag{9}$$

where $g(\cdot)$ is alternately the identity function (in which case estimation is by OLS) or, reflecting that I_t is binary, the inverse logit function[14] (in which cases estimation is by MLE, and η_t is modelled as binomially distributed), and the vector $\boldsymbol{Z}_t$ represents a subset of the information set. The only restriction on the error term η_t is that it is mean zero, hence we employ Newey-West HAC-robust standard errors using the lag selection algorithm of Newey and West (1994).

We follow PT in testing on parsimonious subsets of $\boldsymbol{Z}_t$. The first is exactly as per the empirical example in their Section 2.3: $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}\}$. Results are shown in Table 5 and Table 6. $\hat{\beta}_1$ is small and insignificant so we fail to reject the RPH in this case. However, if we add the sign of the perception error two periods ago (rather than one period ago , which is observable by period t, to get $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}, I_{t-2}\}$ (where we use I_{t-2} rather than I_{t-1} as used by PT because our maintained assumption on the information set does not guarantee that it contains I_{t-1}) then $\hat{\beta}_2$ is significant and RPH is strongly rejected, as seen in Table 7 and Table 8. This rejection is robust to all of linear, logit and probit[15] regression, and of course there are many other covariates we could have tried which may also have given a rejection of RPH. Note that the endogeneity of the lagged dependent variable in the presence of MA(1) errors anticipated under the null, will bias parameter estimates in general, but not under H_0, so rejections of H_0 are not compromised.

Finally, PT proves that the loss function and DGP restrictions they impose are maximally general in the following sense. If the loss function is non-homogeneous then allowing for conditional variance dynamics makes it difficult to obtain testable restrictions that are robust to the loss function. If there are dynamics in higher moments of the DGP then more information on the shape of the loss function is required to obtain testable restrictions, even if the distribution of the forecast error is known.

[14] The probit model gave very similar results to the logit model so is not reported for the sake of space.
[15] Results not reported since they were very similar to logit results.

Table 5: Quantile tests for forecast optimality of inflation perceptions in Sweden, based on equation (9) with $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}\}$ (excluding intervention dummies)

Survey measure →		Including extreme responses						Excluding extreme responses					
Sample →		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T →		180		108		72		180		108		72	
HAC kernel lags		22		12		11		22		11		11	
Linear regression (OLS)													
Constant	$\hat{\alpha}$	0.759***	(0.136)	0.785***	(0.192)	0.787**	(0.314)	0.901***	(0.153)	0.927***	(0.200)	0.689*	(0.367)
			[0.000]		[0.000]		[0.015]		[0.000]		[0.000]		[0.065]
Perceptions,	$\hat{\beta}_1$	-0.014	(0.055)	-0.075	(0.114)	0.028	(0.123)	-0.153	(0.096)	-0.204	(0.139)	0.019	(0.167)
$\tilde{\pi}_t^{12}$			[0.804]		[0.516]		[0.821]		[0.112]		[0.145]		[0.908]
Link test (p-val)		* [0.063]		*** [0.000]		[0.745]		[0.371]		[0.142]		[0.623]	
Logit regression (MLE)													
Constant	$\hat{\alpha}$	1.142*	(0.706)	1.221	(0.889)	1.250	(2.278)	1.771**	(0.781)	1.878*	(1.045)	0.787	(1.777)
			[0.106]		[0.170]		[0.583]		[0.023]		[0.072]		[0.658]
Perceptions,	$\hat{\beta}_1$	-0.069	(0.278)	-0.326	(0.494)	0.220	(0.925)	-0.684	(0.444)	-0.897	(0.676)	0.097	(0.817)
$\tilde{\pi}_t^{12}$			[0.803]		[0.509]		[0.812]		[0.123]		[0.185]		[0.906]
Link test (p-val)		* [0.065]		* [0.036]		[0.727]		[0.244]		[0.137]		[0.635]	

Notes: Newey-West standard errors reported. 'HAC kernel lag' is the truncation lag selected according to Newey and West (1994). The misspecification 'link test' of Pregibon (1979) is a test of the null that the square of the fitted values of the index (the right-hand side of equation (9)) is insignificant in a subsequent regression of $g(I_{t-2}^{12})$ on a constant, those fitted values and their square.

Table 6: Quantile tests for forecast optimality of inflation perceptions in Sweden, based on on equation (9) with $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}\}$ (including intervention dummies)

Survey measure →		Including extreme responses						Excluding extreme responses					
Sample →		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T →		180		108		72		180		108		72	
HAC kernel lag		5		20		11		6		20		11	
Linear regression (OLS)													
Constant	$\hat{\alpha}$	0.798***	(0.114)	0.831***	(0.137)	0.787**	(0.314)	0.809	(0.152)	0.856***	(0.147)	0.689*	(0.367)
			[0.000]		[0.000]		[0.015]		[0.000]		[0.000]		[0.065]
Perceptions,	$\hat{\beta}_1$	-0.034	(0.066)	-0.054	(0.087)	0.028	(0.123)	-0.092	(0.101)	-0.123	(0.106)	0.019	(0.167)
$\tilde{\pi}_t^{12}$			[0.600]		[0.534]		[0.821]		[0.360]		[0.247]		[0.908]
Link test (p-val)		[0.994]		[0.865]		[0.745]		[0.893]		[0.748]		[0.623]	
Logit regression (MLE)													
Constant	$\hat{\alpha}$	1.378**	(0.651)	1.525**	(0.736)	1.250	(2.278)	1.366*	(0.755)	1.584**	(0.743)	0.787	(1.777)
			[0.034]		[0.038]		[0.583]		[0.070]		[0.033]		[0.658]
Perceptions,	$\hat{\beta}_1$	-0.194	(0.359)	-0.277	(0.421)	0.220	(0.925)	-0.428	(0.462)	-0.563	(0.484)	0.097	(0.817)
$\tilde{\pi}_t^{12}$			[0.590]		[0.511]		[0.812]		[0.355]		[0.245]		[0.906]
Link test (p-val)		[0.872]		[0.733]		[0.727]		[0.647]		[0.498]		[0.635]	

Notes: parameter estimates on dummies not reported. See also notes to Table 6.

Table 7: Quantile tests for forecast optimality of inflation perceptions in Sweden, based on equation (9) with $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}, I_{t-2}\}$ (excluding intervention dummies)

Survey measure →		Including extreme responses						Excluding extreme responses					
Sample →		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T →		178		106		72		178		106		72	
HAC kernel lags		22		13		8		22		3		11	
Linear regression (OLS)													
Constant	$\hat{\alpha}$	0.405***	(0.150)	0.462**	(0.186)	0.407	(0.368)	0.415***	(0.135)	0.352**	(0.178)	0.419	(0.374)
			[0.008]		[0.015]		[0.272]		[0.002]		[0.050]		[0.266]
Perceptions,	$\hat{\beta}_1$	0.025	(0.040)	-0.017	(0.073)	0.060	(0.114)	-0.030	(0.065)	-0.045	(0.081)	0.069	(0.154)
$\tilde{\pi}_t^{12}$			[0.539]		[0.818]		[0.601]		[0.639]		[0.582]		[0.655]
LDV, I_{t-2}	$\hat{\beta}_2$	0.390***	(0.113)	0.353**	(0.150)	0.369**	(0.176)	0.441***	(0.082)	0.536***	(0.110)	0.253	(0.165)
			[0.001]		[0.020]		[0.039]		[0.000]		[0.000]		[0.129]
Residual serial corr. (p-val)		*** [0.001]		** [0.014]		*** [0.008]		*** [0.000]		*** [0.001]		*** [0.008]	
Link test (p-val)		[0.712]		[0.860]		** [0.019]		[0.830]		[0.415]		*[0.054]	
Wald test of $\widehat{\boldsymbol{\beta}} = \mathbf{0}$ (p-val)		*** [0.003]		* [0.066]		[0.101]		*** [0.000]		*** [0.000]		[0.313]	
Logit regression (MLE)													
Constant	$\hat{\alpha}$	-0.514	(0.764)	-0.108	(0.872)	-1.046	(2.536)	-0.260	(0.713)	-0.499	(0.969)	-0.497	(1.753)
			[0.501]		[0.902]		[0.680]		[0.716]		[0.607]		[0.777]
Perceptions,	$\hat{\beta}_1$	0.164	(0.263)	-0.090	(0.380)	0.534	(0.985)	-0.171	(0.359)	-0.269	(0.464)	0.369	(0.777)
$\tilde{\pi}_t^{12}$			[0.533]		[0.812]		[0.588]		[0.633]		[0.563]		[0.635]
LDV, I_{t-2}	$\hat{\beta}_2$	1.889***	(0.510)	1.568**	(0.647)	2.715**	(0.891)	1.972***	(0.400)	2.445***	(0.587)	1.201	(0.736)
			[0.000]		[0.015]		[0.015]		[0.000]		[0.000]		[0.103]
Link test (p-val)		[0.838]		[0.919]		* [0.069]		[0.795]		[0.560]		** [0.023]	
Wald test of $\widehat{\boldsymbol{\beta}} = \mathbf{0}$ (p-val)		*** [0.001]		* [0.052]		** [0.027]		*** [0.000]		*** [0.000]		[0.264]	

Notes: The residual serial correlation test has the null of no residual series correlation at up to 12 lags. 'Residual serial corr.' is the Cumby and Huizinga (1992) test of null of no residual serial correlation at up to 12 lags, allowing for endogenous explanatory variables and heteroskedastic errors. The misspecification 'link test' of Pregibon (1979) is a test of the null that the square of the fitted values of the index (the right-hand side of equation (9)) is insignificant in a subsequent regression of $g(I_{t-2})$ on a constant, those fitted values and their square.

Table 8: Quantile tests for forecast optimality of inflation perceptions in Sweden, based on equation (9) with $\boldsymbol{Z}_t = \{\tilde{\pi}_t^{12}, I_{t-2}\}$ (including intervention dummies)

Survey measure →		Including extreme responses						Excluding extreme responses					
Sample →		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12		1993m1–2007m12		1993m1–2001m12		2002m1–2007m12	
Observations, T →		178		106		72		178		106		72	
HAC kernel lags		17		18		8		17		20		11	
Linear regression (OLS)													
Constant	$\hat{\alpha}$	0.571***	(0.156) [0.000]	0.666***	(0.165) [0.000]	0.407	(0.368) [0.272]	0.434***	(0.153) [0.005]	0.387**	(0.170) [0.024]	0.419	(0.374) [0.266]
Perceptions, $\tilde{\pi}_t^{12}$	$\hat{\beta}_1$	-0.006	(0.052) [0.907]	-0.032	(0.071) [0.650]	0.060	(0.114) [0.601]	-0.013	(0.071) [0.854]	-0.023	(0.082) [0.781]	0.069	(0.154) [0.655]
LDV, I_{t-2}	$\hat{\beta}_2$	0.242**	(0.115) [0.036]	0.172	(0.139) [0.216]	0.369**	(0.176) [0.039]	0.379***	(0.087) [0.000]	0.471	(0.079) [0.000]	0.253	(0.165) [0.129]
Residual serial corr. (p-val)		*** [0.000]		** [0.013]		*** [0.008]		*** [0.000]		*** [0.000]		*** [0.008]	
Link test (p-val)		[0.921]		[0.983]		** [0.019]		[0.870]		[0.958]		*[0.054]	
Wald test of $\widehat{\boldsymbol{\beta}} = \mathbf{0}$ (p-val)		* [0.099]		[0.455]		[0.101]		*** [0.000]		*** [0.000]		[0.313]	
Logit regression (MLE)													
Constant	$\hat{\alpha}$	0.239	(0.802) [0.765]	0.770	(0.788) [0.329]	-1.046	(2.536) [0.680]	-0.260	(0.772) [0.737]	-0.385	(0.922) [0.676]	-0.497	(1.753) [0.777]
Perceptions, $\tilde{\pi}_t^{12}$	$\hat{\beta}_1$	-0.031	(0.317) [0.922]	-0.176	(0.367) [0.632]	0.534	(0.985) [0.588]	-0.068	(0.378) [0.858]	-0.137	(0.468) [0.770]	0.369	(0.777) [0.635]
LDV, I_{t-2}	$\hat{\beta}_2$	1.250**	(0.523) [0.017]	0.828	(0.577) [0.151]	2.715**	(0.891) [0.015]	1.725***	(0.407) [0.000]	2.145***	(0.354) [0.000]	1.201	(0.736) [0.103]
Link test (p-val)		[0.187]		[0.784]		* [0.069]		[0.309]		[0.690]		** [0.023]	
Wald test of $\widehat{\boldsymbol{\beta}} = \mathbf{0}$ (p-val)		** [0.050]		[0.349]		** [0.027]		*** [0.000]		*** [0.000]		[0.264]	

Note: parameter estimates on dummies not reported. See also notes to Table 7.

5 Conclusions

Overall there is statistically significant evidence of both mean and median bias. The point estimates of the mean bias – on the order of +½ a percentage point – are substantial compared to relevant benchmarks including the inflation target of 2%±1%, the root mean-squared perception error of1.09%, and then mean and standard deviation of inflation at 1.50% and 1.32% respectively. Thus the bias is also economically significant. If inflation is I(1) then the Mincer-Zarnowitz regression can also be interpreted as showing that this biased persists even in the long-run equilibrium between inflation and perceptions.

This robust bias finding is a new result for Sweden as far as we are aware. It is consistent with findings on US quantitative data (M. F. Bryan & Venkatu, 2001a, 2001b), but contrary to Lein and Maag (2011) which estimates a bias of 0.01% on the 1996m1-2007m8 subsample of our dataset. It is unclear how they get this result since both formally and even by eye from Figure 1 there is substantial positive bias over this period. As in the present paper, Dräger (2014) rejected median-unbiasedness using the Wilcoxon rank test, but she only considered a subsample of our dataset, 1998m1-2007m12, assumed symmetrically distributed perception errors (which our sign test does not) and did not test mean-unbiasedness.

On an earlier sample, 1978-1985, Jonung and Laidler (1988) did not reject mean-unbiasedness. However, their point estimate of bias in the Mincer-Zarnowitz regression is larger than ours and the non-rejection may be the result of low power due to a shorter and lower frequency sample than ours. Applying the sign test to their data we reject the null of median-unbiasedness with a p-value of 0.0009[16], which illustrates the added value of the nonparametric tests compared to the parametric tests.

By contrast, our finding of strong and significant serial correlation in perception errors beyond the first lag is consistent with the aforementioned Swedish literature.

The above results imply rejection of RPH under MSE and MAE loss, confirming the finding of earlier literature on earlier samples and subsamples of this Swedish dataset. However, our rejection of RPH under the more general combinations of restrictions on the (unknown) loss function and DGP covered by the indicator test of Patton and Timmermann (Patton & Timmermann, 2007) is new in the literature as far as we are aware. The combinations covered by this rejection include any loss function that is solely a function of perception error in combination with any inflation dynamics are limited to the conditional mean and variance.

Of course it is possible that there are dynamics in higher moments of inflation and/or that the loss function depends on variables other than perception error (e.g. the level of inflation). For example, non-linearities in the DGP when inflation is close to zero or very high (and thus at greater risk of a deflationary or hyperinflationary spiral, though this is unlikely to have been a major issue in our sample) might induce conditional skew dynamics. Similarly, the loss from under-predicting inflation

---

[16] The raw data in their Chart I shows that perception errors were positive in 4 out of their 25 observations, and the positive errors were similar in magnitude to the negative errors. The sign statistic $S_{odd} + S_{even}$ (appropriate under the null of zero median and independent, rather than MA(1), errors) is distributed as Binomial(25, 0.5) under the null, so the two-sided p-value for 4 'successes' is 0.0009.

may be greater when inflation is high and threatening to de-anchor from the official target, and the loss from over-predicting inflation may be greater when it is low and deflation threatens. From a consumer perspective the loss could be micro-founded in terms of the opportunity cost of failing to take appropriate action to optimise real income, real balance sheet position, or to insure against indirect effects such as loss of employment that may become more probable when inflation is far from target.

The other maintained assumptions might alternatively explain our rejections of RPH without having to reject the assumption of fully optimising behaviour.

Statistical learning rules relax the assumption that the consumer knows the data generating process *a priori* (see Evans & Honkapohja, 2001 for an overview). This seems unlikely to offer a complete explanation since the rate of learning would have to be very slow to for bias to persist in our 15-year span of data during which there does not appear to have been a major regime shift.

Relaxing the assumption that the reference index is CPI for all consumers offers a perhaps more plausible route to rationalising perceptions. For example, consumer inflation perceptions may be disproportionately affected by the prices of particularly salient or frequently purchased goods such as petrol, bread and milk (e.g. Brachinger, 2008; Curto Millet, 2006; Ehrmann, 2006). Some recent evidence is consistent with this. Curto Millet (2006) was able to rationalise a subsample of our perceptions data in the long-run equilibrium of a cointegrated VECM by allowing different weights on 12 sub-indices of CPI inflation. Eckley (2009) found seasonality in inflation perceptions that is not present in actual inflation. This might be rationalised by seasonally varying consumption patterns.

The empirical challenge ahead lies in testing RPH under further relaxation of the assumptions on the reference index, information set, conditional distribution and prior set of loss functions, to establish what combinations of assumptions (if any) can rationalise consumer inflation perceptions. Extending tests of RPH to the plausible DGP/loss function combinations mentioned above, using the flexible spline approximation to the loss function proposed by Patton & Timmermann (2007), could be one next step.

Whether or not perceptions are fully rationalisable, they at least *partially* incorporate the available information as reflected in $\hat{\beta}$ close to unity in the Mincer-Zarnowitz regressions and the fact that root-mean square error (RMSE) of inflation perceptions (0.679 percentage points) is substantially smaller than RMSE of 12-month ahead inflation expectations from the same survey and period (1.143 percentage points)[17].

6 Implications

Our findings have important implications for wider theoretical and empirical work in at least three areas.

6.1 Evaluating theory models of inflation perceptions and expectations

A positive bias is inconsistent with the near-rationality hypothesis of Akerlof, Dickens, & Perry (2000) applied to inflation perceptions considered as nowcasts. One version of their model predicts that

[17] We exclude 1993 from this comparison for the same reason that we dummy 1993 in our regressions.

members of the labour force have rational expectations (and thus implicitly rational perceptions). The other version predicts that the only systematic expectation errors members of the labour force[18] would make are negative. This adds support to the finding of Bryan & Palmqvist (2005) of little support for near-rationality on forward-looking expectations from the same survey.

The finding of bias even in the *long-run* equilibrium (assuming inflation is I(1) and cointegrated with perceptions) is further inconsistent with several popular models that seek to rationalise REH/RPH violations by relaxing one or two of the bundled assumptions, but that still that perceptions tend to equality with inflation in the long-run. These include rational inattention (Reis, 2006; Sims, 2003); sticky information (Mankiw & Reis, 2002); and the epidemiological model of Carroll (2003).

6.2 Macroeconomic modelling

Consumer inflation perceptions play a major role in most contemporary macroeconomic models via their involvement in decisions about durables purchases, saving, debt management and wage negotiations. It is assumed that agents observe nominal variables but care about real variables, and adjust between the two using perceived inflation. Perception errors are typically assumed (usually implicitly) to be deterministically zero or if stochastic then to be symmetrically distributed with mean zero. Our finding of perceptions bias implies a wedge between actual and perceived values of real variables which may represent a substantial modification to some models, and our finding of serially correlated perceptions errors would constitute an independent source of persistence.

Inflation perceptions also play an important indirect role in macroeconomic models through their influence on forward-looking inflation expectations, where these are modelled as adaptive (extrapolating perceptions of past and/or current inflation), forward-looking (in which case the forecast model is likely to be conditioned on inflation perceptions), or some hybrid of the two. Empirically, consumer inflation perceptions are so consistently close to consumer inflation expectations in our Swedish data (see Figure 2 in the Appendix) that a model of inflation expectations as extrapolating currently perceived inflation into the next period seems plausible. Similarly, Simmons and Weiserbs (1992) find that predicting inflation using survey expectations relative to the base of perceived (rather than actual) inflation “works rather well”, and Benford & Driver (2008) find that 40% of UK households claim their current perception of inflation is a very important factor in setting their expectations. Reviewing the literature, Raynard et al (2008) conclude that the available evidence points to projecting current inflation into the future to be a plausible model of people’s actual cognitive heuristics. Therefore inflation perception errors are likely to feed through to forward-looking inflation expectation errors, thus violating the assumption of REH made in many macroeconomic models.

6.3 Quantification of forward-looking inflation expectations

Most of survey data on forward-looking inflation expectations derives from responses to *qualitative* questions like this one from European Commission Directorate-General for Economic and Financial Affairs (2014): “By comparison with the past 12 months, how do you expect that consumer prices

[18] Strictly our perceptions measure is representative of the population at large rather than the labour force. Palmqvist and Strömberg (2004) reports that for 2001m11 to 2004m5 that perceptions of the 63% of the population in the labour force exhibits smaller bias than the population at large, but the bias is still substantial and positive.

will develop in the next 12 months? Increase more rapidly / increase at the same rate / increase at a slower rate / stay about the same / fall / don't know". Typically a *quantitative* inflation expectation value is desired, whether to study its properties as a cardinal forecast of official inflation figures, as conditioning data for empirical macroeconomic forecast models, or as an input to policy-making.

Producing quantitative estimates from the available qualitative data necessitates assumptions about the properties of the relationship between inflation perceptions and outturns or, equivalently, about the properties of inflation perception errors. The 'probability' method of quantification Carlson & Parkin (1975) for trichotomous response data assumes that inflation perceptions are error-free. Its widely-used extension to pentachotomous data due to Batchelor and Orr (1988) requires a scaling series upon which inflation perceptions are assumed to be based. Defining such a series implies restrictions on inflation perceptions errors. One restriction commonly imposed is that inflation perceptions are unbiased (e.g. Berk, 1999; Forsells & Kenny, 2004; Gerberding, 2001; Lein & Maag, 2011). Less stringently, Dräger, Menz, & Fritsche (2014) assume that households correctly perceive a medium-term trend of the true inflation rate extracted using a Hodrick-Prescott filter, which would imply long-run unbiasedness of inflation perceptions.

Our findings of inflation perceptions bias in both short and long run thus raise questions over the quality of the derived quantitative consumer/household inflation expectations numbers that are widely used in monetary policy making (e.g. Bernanke, 2004).

References

Akerlof, G. A., Dickens, W. T., & Perry, G. L. (2000). Near-Rational Wage and Price Setting and the Long-Run Phillips Curve. *Brookings Papers on Economic Activity*, *31*(1), 1–60.

Altissimo, F., Ehrmann, M., & Smets, F. (2006). Inflation Persistence and Price-Setting Behaviour in the Euro Area. *ECB Occasional Paper Series*, *46*.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, 27–41.

Batchelor, R. A., & Orr, A. B. (1988). Inflation Expectations Revisited. *Economica*, *55*(219), 317–331.

Benford, J., & Driver, R. (2008). Public attitudes to inflation and interest rates. *Bank of England Quarterly Bulletin*, *Q2*, 148–157.

Berg, C. (2000). *Inflation Forecast Targeting: The Swedish Experience* (No. 100) (pp. 1–35). Retrieved from http://hdl.handle.net/10419/82411

Berk, J. M. (1999). Measuring inflation expectations: a survey data approach. *Applied Economics*, *31*(11), 1467–1480. doi:10.1080/000368499323337

Bernanke, B. S. (2004). Remarks by Governor Ben S. Bernanke. *The Logic of Monetary Policy*.

Brachinger, H. W. (2008). A new index of perceived inflation: Assumptions, method, and application to Germany. *Journal of Economic Psychology*, *29*(4), 433–457.

Bruine de Bruin, W., van der Klaauw, W., Topa, G., Downs, J. S., Fischhoff, B., & Armantier, O. (2012). The effect of question wording on consumers' reported inflation expectations. *Journal of Economic Psychology*, *33*(4), 749–757. doi:10.1016/j.joep.2012.02.001

Bryan, M. F., & Venkatu, G. (2001a). *The curiously different inflation perspectives of men and women*. Federal Reserve Bank of Cleveland, Research Dept.

Bryan, M. F., & Venkatu, G. (2001b). *The demographics of inflation opinion surveys*. Federal Reserve Bank of Cleveland, Research Dept.

Bryan, M. R., & Palmqvist, S. (2005). *Testing near-rationality using detailed survey data*. *Sveriges Riksbank Working Paper Series*. Stockholm: Sveriges Riksbank .

Campbell, B., & Ghysels, E. (1995). Federal Budget Projections: a a nonparametric assessment of bias and efficiency. *The Review of Economics and Statistics*, *77*(1), 17–31.

Carlson, J. A., & Parkin, M. (1975). Inflation Expectations. *Economica*, *42*(May), 123–138.

Carroll, C. D. (2003). Macroeconomic Expectations of Households and Professional Forecasters. *Quarterly Journal of Economics*, *118*(1), 269–298.

Centralbyrån, S. (2001). The Swedish consumer price index: a handbook of methods. Stockholm: Statistics Sweden (SCB).

Clements, P., & Hendry, D. F. (1998). *Forecasting Economic Time Series*. *The Marshall Lectures on Economic Forecasting*. Cambridge, UK: Cambridge University Press.

Cumby, R. E., & Huizinga, J. (1992). Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions. *Econometrica*, *60*(1), 185–195. Retrieved from http://www.jstor.org/stable/2951684

Curto Millet, F. (2006). *Inflation Expectations, Labour Markets and EMU*. *Balliol College*. University of Oxford, Oxford.

Dias, F., Duarte, C., & Rua, A. (2010). Inflation expectations in the euro area: are consumers rational? *Review of World Economics*, *146*(3), 591–607. doi:10.1007/s10290-010-0058-6

Doornik, J., & Hansen, H. (2008). An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics*, *70*, 927–939. doi:10.1111/j.1468-0084.2008.00537.x

Dräger, L. (2014). Inflation perceptions and expectations in Sweden - Are media reports the missing link? *Oxford Bulletin of Economics and Statistics*, (2006), n/a–n/a. doi:10.1111/obes.12078

Dräger, L., Menz, J.-O., & Fritsche, U. (2014). Perceived inflation under loss aversion. *Applied Economics*, *46*(3), 282–293.

Dufour, J.-M. (1981). Rank tests for serial dependence. *Journal of Time S*, *2*(3), 117–128.

Eckley, P. (2009). *Consumer inflation perceptions: (non)rationality, seasonality, and the link to forward-looking expectations*. University of Oxford. Retrieved from http://solo.bodleian.ox.ac.uk/primo_library/libweb/action/dlDisplay.do?vid=OXVU1&docId=oxfaleph017160051

Ehrmann, M. (2006). Rational inattention, inflation developments and perceptions after the euro cash changeover. *ECB Working Paper Series*, *588*.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). The Econometric Society. *Econometrica*, *64*(4), 813–836.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1007.

European Commission Directorate-General for Economic and Financial Affairs. (2014). *The Joint Harmonised EU Program of Business and Consumer Surveys: User Guide* (pp. 1–49). Brussels. Retrieved from http://ec.europa.eu/economy_finance/db_indicators/surveys/documents/bcs_user_guide_en.pdf

Evans, G. W., & Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Forsells, M., & Kenny, G. (2004). Survey Expectations, Rationality and the Dynamics of Euro Area Inflation. *Journal of Business Cycle Measurement and Analysis*, *3*(1), 13–41. doi:10.1787/jbcma-v2004-art3-en

Gerberding, C. (2001). The information content of survey data on expected price developments for monetary policy. *Deutsche Bundesbank ERC Discussion Papers*.

Granger, C. W. J., & Newbold, P. (1986). *Forecasting Economic Time Series, 2nd edition*. Academic Press.

Hansen, L. P., & Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, *88*, 829–853.

Holden, K., & Peel, D. A. (1990). On testing for unbiasedness and efficiency of forecasts. *Manchester School*, *58*, 120–127.

Jonung, L., & Laidler, D. (1988). Are Perceptions of Inflation Rational? Some Evidence for Sweden. *The American Economic Review*, *78*(3), 1080–1087.

Konjunktur Institutet. (n.d.). User guide to the Economic Tendency Survey. Retrieved January 05, 2015, from http://www.konj.se/download/18.2cabf50a141002857ee147b/User-Guide-Economic-Tendency-Survey.pdf

Lein, S. M., & Maag, T. (2011). The Formation of Inflation Perceptions - Some Empirical Facts for European Countries. *Scottish Journal of Political Economy*, *58*(2), 155–188.

Mankiw, N. G., & Reis, R. (2002). Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, *117*(4), 1295–1328.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. A. Mincer (Ed.), *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (Vol. I, pp. 1–46). New York: National Bureau of Economic Research. Retrieved from http://www.nber.org/chapters/c1214

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*(3), 315–335.

Newey, W., & West, K. D. (1987). A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, (55), 703–708.

Newey, W., & West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, *61*, 631–653.

Palmqvist, S., & Strömberg, L. (2004). Households Inflation Opinions - A Tale of Two Surveys. *Sveriges Riksbank Economic Review*, *4*, 23–42.

Patton, A. J., & Timmermann, A. (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, *102*(480), 1172–1184. doi:10.1198/016214506000001176

Pregibon, D. (1979). *Data Analytic Methods for Generalized Linear Models*. Thesis (Ph.D.)--University of Toronto. Retrieved from http://www.worldcat.org/title/data-analytic-methods-for-generalized-linear-models/oclc/640147904

Ranyard, R., Missier, F. Del, Bonini, N., Duxbury, D., & Summers, B. (2008). Perceptions and expectations of price changes and inflation: A review and conceptual framework. *Journal of Economic Psychology*, *29*(4), 378–400. doi:10.1016/j.joep.2008.07.002

Reis, R. (2006). Inattentive consumers. *Journal of Monetary Economics*, *53*(8), 1761–1800.

Simmons, P., & Weiserbs, D. (1992). Consumer Price Perceptions and Expectations. *Oxford Economic Papers*, *44*(1), 35–50.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.

Wallis, K. F. (1986). Forecasting with an Econometric Model: The “Ragged Edge” Problem. *Journal of Forecasting*, *5*(December 1985), 1–13.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, *1*(6), 80–83.

Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters*, (21), 45–48.

Appendix

Monetary policy timeline

Table 9: Monetary policy events in Sweden, 1991–2007

Date	Monetary policy event
1991m5	Krona unilaterally pegged to the European Currency Unit (ECU)
1992m11	Fixed exchange rate regime collapses on 19 Nov; krona de-pegged from ECU and trade-weighted exchange rate rapidly depreciates, losing about 25% of its value by end of 1993. This induces an inflationary impulse through 1993 and 1994
1993m1	Announcement (on 14 Jan) of plan to begin inflation targeting in 1995 with annual increase in CPI to be limited to of 2%±1%
1993–1994	Riksbank aims monetary policy at "preventing the inflationary impulse, due to the depreciation of the krona and changes in indirect taxes, from causing an increase in the underlying rate of inflation"*
1995m1	Inflation target becomes officially operational
1995m1–1995m12	Implicit inflation forecast targeting (not based on published inflation forecasts)
1996m1–1997m12	Explicit inflation forecast targeting (based on published inflation forecasts)
1998m1–2007m12	Inflation distribution forecast targeting

Notes: Based on (Berg, 2000). * Source: Sveriges Riksbank, Press Release No. 5, 1993.

Unit root tests

Table 10: Unit root tests

Variable	1993m1-2007m12		1994m1-2007m12	
	Test statistic	95% critical value	Test statistic	95% critical value
π_t^{12}	-0.966		-1.732	
$\tilde{\pi}_t^{12}$ including extreme responses	-2.578	-2.046	-2.555	-2.053
$\tilde{\pi}_t^{12}$ excluding extreme responses	-1.631		-2.632	

Notes: Test statistic is the DF-GLS test statistic of Elliott et al. (1996) using one lag as selected by Schwarz information criterion, d93 and dgfk dummies but no deterministic trend (consistent with visual examination of the data).

Inflation expectations cf. perceptions

The expectations question in the SCB's monthly Consumer Tendency Survey is worded:

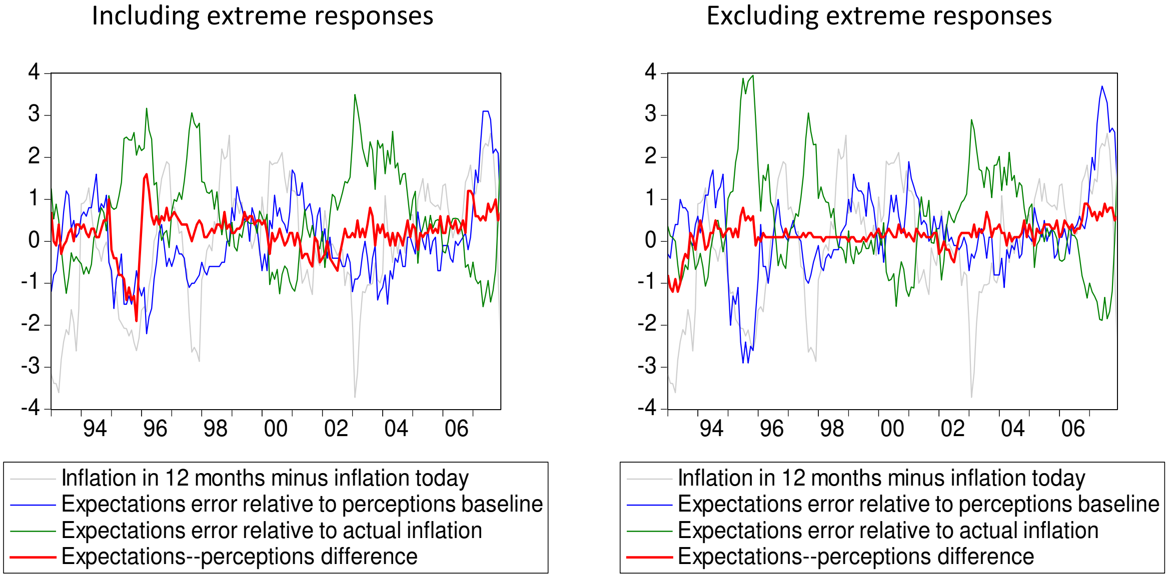
"Compared with today, by what percentage do you think that prices will go up (i.e. the rate of inflation 12 months from now)?"

Figure 2 shows four series:

- expectations error relative to actual inflation: $\pi_{t|t+12}^{e,12} - \pi_{t+12}^{12}$
- expectations error relative to perceptions baseline $\pi_{t|t+12}^{e,12} - \tilde{\pi}_{t+12}^{12}$
- expectations–perceptions difference: $\pi_{t|t+12}^{e,12} - \tilde{\pi}_t^{12}$
- 12-month change in inflation: $\pi_{t+12}^{12} - \pi_t^{12}$

where $\pi_{t|t+12}^{e,12}$ is the expectation formed in period t about π_{t+12}^{12}.

Figure 2: Difference between expectations of inflation over next 12 months, and perceptions of inflation over last 12 months, based on responses from same survey, Sweden, 1993m1–2007m12



Note: SCB/GfK did not publish their criterion for classifying a response as 'extreme', so we do no know whether the extreme expectations responses are from the same respondents as the extreme perceptions responses. In more detailed (and rigorous) econometric analysis we would clarify this.

Notice that the expectations–perceptions difference (today's forecast of future inflation cf. today’s perception of today’s inflation) is close to mean-zero (mean of 0.21 or 0.17 including and excluding extreme responses respectively) and is considerably less volatile than the other series plotted (standard deviation of 0.49 or 0.33 compared to 0.95 or 1.05 for the expectations error relative to perception baseline, 1.14 or 1.21 for expectation errors relative to actual inflation, and 1.44 for 12-month change in inflation).

This strongly suggests that inflation expectations are based more on current inflation perceptions than on actual inflation, in which case many of the properties of inflation expectations examined in the literature to date may be inherited from the perceptions formation mechanism.