

# Time Variation of Regression Coefficients related to Macroeconomic News affecting Currency Prices

Victor Mitchell

November 5, 2019

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data</b>	<b>2</b>
<b>3</b>	<b>Building blocks of the models</b>	<b>2</b>
3.1	The Stable Linear Model that is Time Invariant . . . . .	2
3.2	The Gaussian Random Walk . . . . .	5
<b>4</b>	<b>Testing for instability of the news impact parameter</b>	<b>5</b>
4.1	The quasi-Local-Level Test . . . . .	5
4.2	CUSUM and CUSUM-squared tests . . . . .	8
<b>5</b>	<b>Parameter Path Estimations</b>	<b>10</b>
5.1	Weighted Average Risk Minimization . . . . .	10
5.2	Standard Recursive Time Variable Parameter Algorithm (STVP) . . . . .	13
<b>6</b>	<b>Discussion</b>	<b>16</b>
6.1	Differences and Irregularities of Two Paths . . . . .	16
6.2	Accuracy of the Estimations . . . . .	16
6.3	One-Step ahead Forecasts . . . . .	16
6.4	Use of the Post-Announcement 5 Minute Window . . . . .	17
<b>7</b>	<b>Appendix</b>	<b>19</b>
7.1	Parameter Paths . . . . .	19
7.2	Stochastic Recursive Least Squares Algorithm (SRLS) . . . . .	19
7.3	Standard Recursive Time Variable Parameter Algorithm (STVP) . . . . .	19
7.4	Steps to obtain qLL statistic . . . . .	27
7.5	Steps to obtain WAR minimization path . . . . .	27
7.6	Variance for credible intervals for the WAR path . . . . .	27
	<b>References</b>	<b>27</b>

abstract here when I'm done with the writing and have all results.

## 1 Introduction

Some macroeconomic figures used to gauge the health of a nation's economy are released to the public on a predetermined schedule. These include for example the Non-Farm Employment change that is released on the first friday of every month informing economists and investors alike of the status of employment in the

United States.

Standard economic theory helps us understand that an increase of interest rates is warranted when economies are performing well and prices are generally increasing. Those who decide to increase national interest rates, the central banks, typically refer to measures of inflation in order to make their decisions. Because of this, investors and traders alike pay close attention to news releases (such as inflation, and also the Non-Farm-Payrolls in the case of the United States) and react according to the results. These news releases are not made public until specific times on specific days and since investors and traders react to the same news the moment it is released, the result is often a violent reaction of price in one direction or another. The common discourse is that the direction and the magnitude of the change of price depends on the difference between the expectation of the market (combined expectation of worldwide investors) and the result of the news release.

In this paper, we decide to use currency pairs to measure the price shocks. As certain news pertaining to a particular country affects its respective currency more than other ones, it makes sense to observe the currency most relevant to the news announcement. As currency prices are typically measured in pairs, the second chosen currency will be another major currency that is known for its high liquidity (EUR, USD or CHF) and does not have any other news announcements at the same time<sup>1</sup>. As an example, we would use the USD/CAD currency pair to measure the effect of the Canadian Consumer Price Index (CPI).

The paper of (Andersen et al. 2003) reveals that over the time period between 1987 and 2002 there has been little time-variation in the reaction to news. Some more recently published literature of (Ben Omrane et al. 2019) involving an analysis on euro-dollar contracts has determined that unlike the the decade(s) encompassing the “Great Moderation” where there was lower relative volatility in the financial markets, the time period between 2004 and 2014 is characterized by evolving reactions to macroeconomic news.

The objective of this paper is two-fold. Firstly, we aim to employ existing methodologies on new data to identify whether or not instability of market reaction is present. When reactions do change over time, we attempt to estimate the path that it takes using different procedures suggested by the literature given that the instability has been identified.

## 2 Data

The minute-by-minute OHLC Data of 7 currency pairs were collected from the Metatrader5 platform between 2008 and the end of 2019. This represents over 4 million data points for each currency pair considered in the study. Only a small fraction of this data is actually used since we consider only the 5 time frame from when each piece of monthly or quarterly news is released until 5 minutes afterwards. In regards to the collection of macroeconomic figures data, the expectation numbers from the ForexFactory website were used<sup>2</sup> The process of combining of the two separate datasets was complex and involved careful consideration of erratic scheduling changes of the organizations releasing the figures as well as the daylight savings and timezones of all considered data.

## 3 Building blocks of the models

### 3.1 The Stable Linear Model that is Time Invariant

Being consistent with previous literature on the subject, the first step involves estimating the impact that each piece of news has on its respective currency assuming 1.) That the news effects are constant over time. 2.) The surprise element  $S_t$  of the regression is evaluated as:

$$S_t = \frac{A_t - E_t}{\sigma_d} \quad (1)$$

---

<sup>1</sup>When simultaneous news cannot be avoided, a sequence of stability tests will be applied to ensure time variation is identified on a specific news release. Alternatively, when simultaneous news for two announcements occur at the same time but only for a fraction of the sample, those observations are omitted (this was the case for the Canadian Retail Sales and Consumer Price Index).

<sup>2</sup>The expectations of most online sources such as “Investing.com” or “DailyFX” are very similar. Unfortunately, the aggregation methods are not disclosed to the public.



Figure 1: Minute-by-minute candlechart of the USD/CAD asset on the 19th of June between 15h00 and 16h00 GMT+2. An example of the sudden price change that can occur during one of the news releases. Green/Red candles represent an increase/decrease in price. The gray indicator is amount of trading activity taking place (a rough guideline as it only includes transactions from that particular brokerage).

Table 1: Summary of the news figures considered in the study

Country	News.Event	Pair.used	GMT.Time	Frequency	Observations	Dates
<b>Single News</b>						
Canada	Consumer Price Index	USD/CAD	13:30	Monthly	103	Jan 2008 to Dec 2019
Canada	Core Retail Sales	USD/CAD	12:30	Monthly	94	Oct 2008 to Dec 2019
United States	Consumer Price Index	USD/CHF	13:30	Monthly	135	Oct 2008 to Dec 2019
New Zealand	Consumer Price Index	NZD/USD	21:45	Quarterly	43	Jan 2009 to Oct 2019
Australia	Consumer Price Index	AUD/USD	00:30	Quarterly	48	Jan 2008 to Oct 2019
Australia	Retail Sales	AUD/USD	00:30	Monthly	135	Oct 2008 to Dec 2019
United Kingdom	Consumer Price Index	GBP/USD	09:30	Monthly	135	Oct 2008 to Dec 2019
United Kingdom	Retail Sales	GBP/USD	09:30	Monthly	135	Oct 2008 to Dec 2019
<b>Grouped News</b>						
United States	Average Hourly Earnings Change	USD/CHF	13:30	Monthly	135	Oct 2008 to Dec 2019
United States	NonFarm Employment Change	USD/CHF	13:30	Monthly	135	Oct 2008 to Dec 2019
United States	Unemployment Rate	USD/CHF	13:30	Monthly	135	Oct 2008 to Dec 2019
Canada	Employment Change	USD/CAD	13:30	Monthly	135	Oct 2008 to Dec 2019
Canada	Unemployment Rate	USD/CAD	13:30	Monthly	135	Oct 2008 to Dec 2019
Australia	Employment Change	AUD/USD	00:30	Monthly	135	Oct 2008 to Dec 2019
Australia	Unemployment Rate	AUD/USD	00:30	Monthly	135	Oct 2008 to Dec 2019

Table 2: Results of regressions tests - HAC standard errors

News.Event	M5.Coefficient	std.error	HAC.std.error
<b>Single News</b>			
UK CPI	12.681***	1.729	2.601
CA CPI	-8.6286***	1.958	2.502
CA CRS	-10.756***	1.624	2.249
US CPI	3.968**	1.207	1.23
NZ CPI	24.109***	2.910	4.727
AU CPI	22.293***	4.145	4.226
AU RET	9.647***	1.215	2.656
UK RET	16.574***	1.968	2.572
<b>Grouped News</b>			
US AHE	10.295***	2.577	2.665
US NFP	17.938***	2.572	4.143
US URÂ°	1.975	2.579	2.198
CA EMC	-25.588***	3.940	4.211
CA UR	1.053	3.940	4.262
AU EMC	21.59571***	1.842	2.780
AU UR	-11.93102***	1.842	1.828

*Note:* The result of OLS estimation, referred to as the 'time invariant' or 'stable' case is presented in this table. The standard errors of the estimator as well as the Newey-West corrected standard errors are also included. The \*, \*\*, \*\*\* are for 10%, 5% and 1% significance levels respectively.

$A_t$  is the actual result of the news at time  $t$ ,  $E_t$  is the expected result aggregated from experts and  $\sigma_d$  is the empirical standard deviation of this difference over the entire sample. Thereafter, we use this surprise element in a first simple OLS model.

$$R_t = \beta_0 + \beta_1 S_t + \varepsilon_t \quad (2)$$

With  $R_t$  as the 5-minute currency return result and  $\varepsilon_t$  is the usual assumed normally distributed error. Moreover, because we are working with a dataset where subsequent observations are suspected to be related to one another, one could expect that the errors of the basic model above be autocorrelated. Specifically, adjacent  $R_t$  would be more similar to one another than reactions that are separated in time to a greater extent. In this case, the inference on the  $\beta_1$  would be flawed. Previous researchers have used what is called Heteroskedasticity and Autocorrelation-Consistent (HAC) estimators for the variance of the OLS estimator  $\beta_1$ . Using the Newey-West estimator for this variance from Newey and West (1987), we use modified standard errors of the  $\beta_1$  in our results. If we are wrong in our assumption in some of the news instances, and there is no underlying autocorrelation of the  $R_t$  observations, our estimation of  $\beta_1$  is less efficient than the original estimator in those cases. Nonetheless, it remains consistent and ensures we avoid type 1 error of rejecting a true null hypothesis suggesting  $\beta_1 = 0$ .

Table 2 shows the result of the different  $\beta_1$  coefficients for separate news reports. The construction of the truncation parameter in the Newey-West estimator is such that our monthly news reports consider 2 autocorrelation coefficients whereas the quarterly ones only contain 1. This is due to the difference in the number of observations in our data. A higher estimated autocorrelation between the errors of the regression will result in a stronger correction of the variance of  $\beta_1$ . In all of the news in Table 2, the higher standard error does not affect the significance of the term. While it is not formal evidence, we suspect that the regressions where the HAC standard error is larger than its unedited counterpart contain unknown regressors

that may come from any of the findings that are well elaborated in other literature such as in Goldberg and Grisse (2013). Overall, these results and that of the Durbin-Watson test show that there is little evidence for significant autocorrelation in the residuals.

$$R_t = \beta_0 + \beta_{1,t}S_t + \varepsilon_t \quad (3)$$

### 3.2 The Gaussian Random Walk

The methods discussed in this paper use Gaussian Random Walks (GRW) extensively. This section serves to underline the features of these processes which will later be essential to the understanding of tests and path estimations. Furthermore, it hopefully helps one fully appreciate the assumptions that will be made when using them and how these processes can be used to help decipher the variation of the  $\beta$ .

The GRW is a sequence of *i.i.d* random variables where one realization at time  $t$  follows the distribution:

$$X_t \sim \mathcal{N}(\mu_t, \sigma_t^2) \quad (4)$$

Using the independence assumption of normally distributed random variables, their sum and the distribution of their sum is then given by 3.

$$U = \sum_{t=1}^T X_t \quad U \sim \mathcal{N}(\mu_t T, \sigma_t^2 T) \quad (5)$$

In theory the path that the  $\beta$  could take could resemble any of the paths in Figure 2. The process depicted by the Y2 is a baseline case where the parameter would not vary in time. The Y3 process is a random walk with a  $\mu$  of 0 and a fixed  $\sigma$  which is in reality the same as white noise. The Y1 is a random walk with a constant negative  $\mu$ . Finally, the Y4 process depicts a scenario with one significant break in the path of  $\beta$  and is the combination of 2 separate random walks with separate constant  $\mu$  and  $\sigma$ . Within the context of our problem, it is important to consider all of these possibilities as the potential evolution in the reactions could be erratic in nature (frequent and haphazard change), or gradual (slow and constant change) over time. We would like to detect the changes regardless of the scenario.

## 4 Testing for instability of the news impact parameter

### 4.1 The quasi-Local-Level Test

There exists many ways to test whether  $\beta_t$  is time dependent or not. We first choose the methodology of the authors of (Elliott and Müller 2006) and replicate their method. The advantage of their test is that it identifies instability no matter whether it comes in the form of a single break, many breaks, or a continuous change (all of which are feasibly possible in the context of our investigation). Furthermore, the result of the Monte Carlo simulations over different processes also showed that compared to other tests that exist, the qLL test has the most power in smaller samples.

The quasi-Local-Level (qLL) test enables one to test for many different types of persistent processes of the  $\beta_t$ . It is explained that many of these breaking processes can have a “temporary memory” (strictly speaking are strongly mixing) but will be well approximated by a Wiener process.<sup>3</sup> This is extremely practical in our scenario as there are many possibilities for the possible variation of the  $\beta_t$ . The Null Hypothesis implies there is a stable parameter as in a familiar OLS regression. We obtain the likelihood under the Null assuming that the  $R_t$  observations are independently and identically distributed (and therefore so are their first differences):

---

<sup>3</sup>Theorem 7.30 of (White 2001) can be applied since certain assumptions are made about the process

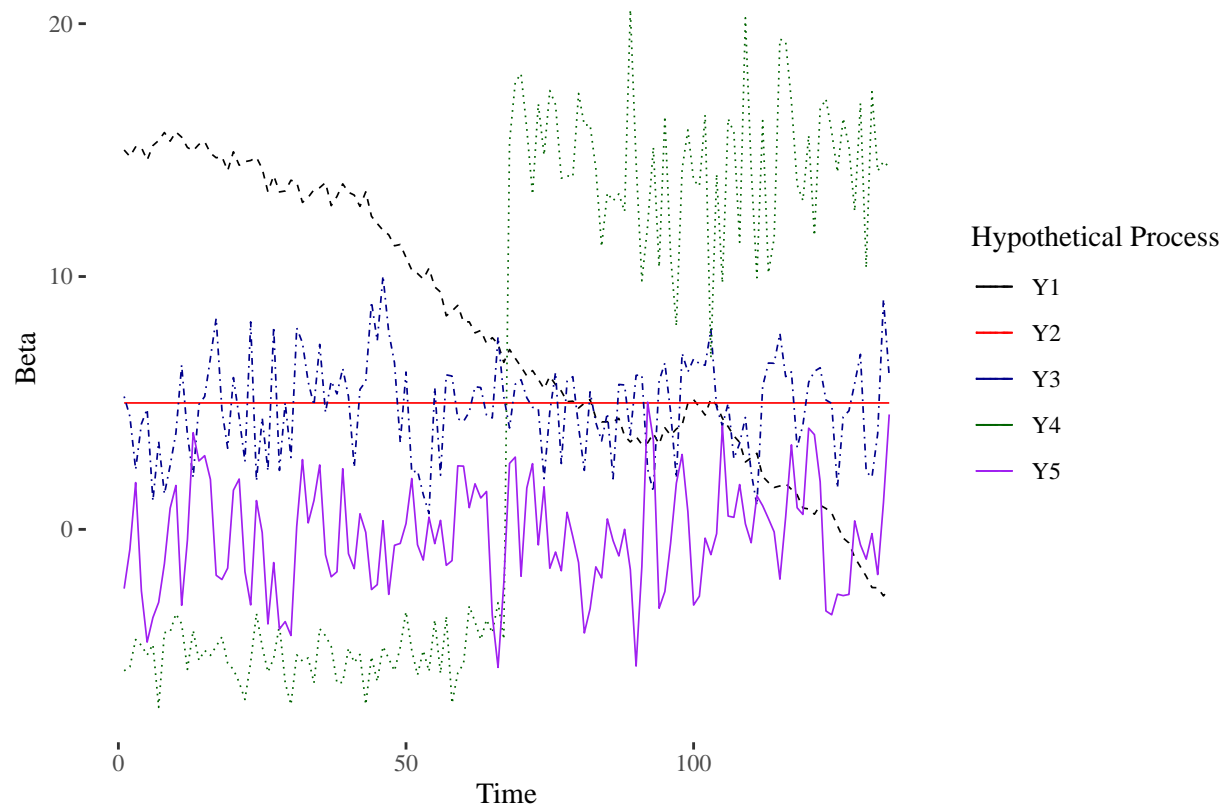


Figure 2: Hypothetical Paths of the Beta

Table 3: Instability Test Results

News.Event	qLL	CUSUM	CUSUM.sq
<b>Single News</b>			
UK CPI	-12.964***	n.s	**
CA CPI	-13.015***	n.s	***
CA CRS	-9.713**	n.s	**
US CPI	-21.582***	n.s	***
NZ CPI	-7.267*	n.s	***
AU CPI	-9.022**	n.s	***
AU RET	-17.623***	**	***
UK RET	-4.503	n.s	***
<b>Grouped News</b>			
<b>US Batch</b>			
Test 1 All News	-28.14***	*	***
Test 2 AHE&NFP	-24.015***		
Test 3 NFP&UR	-20.326***		
Test 4 AHE&UR	-7.078		
Test 5 NFP	-17.643***		
<b>CA Batch</b>			
Test 1 All News	-9.30	n.s	***
<b>AUD Batch</b>			
Test 1 All News	-22.912***	n.s	***
Test 2 EMC	-6.85		
Test 3 UR	-5.66		

*Note:* Results of the three instability tests performed for each piece of macroeconomic news. Multiple tests were performed for some news that occurred simultaneously because the tests are able to detect instability but not pinpoint the exact source of instability for more than 1 figure. The n.s, \*, \*\*, \*\*\* are for non-significant, 10%, 5% and 1% levels of significance respectively.

Table 4: Asymptotic Critical Values of the qLL Statistic

k	1	2	3	4	5
1%	-11.05	-17.57	-23.42	-29.18	-35.09
5%	-8.36	-14.32	-19.84	-25.28	-30.60
10%	-7.14	-12.80	-18.07	-23.37	-28.55

*Note:* Extract of the critical values of the qLL Statistic. k represents the number of potential unstable coefficients (number of parameters in the model) whereas 1%, 5% and 10% are the significance levels where a lower value is stronger evidence that instability is present.

$$L_{H0}(\beta_0, \beta_1, \sigma^2) = \log \prod_{t=1}^T p(\Delta R_t | S_t; \beta_0, \beta_1, \sigma^2) \quad (6)$$

$$= -\frac{T}{2} \log(2\pi) - T \log(\sigma) - \frac{1}{2\sigma^2} \sum_{t=1}^T (\Delta R_t - (\Delta\beta_0 + \Delta\beta_1 S_t))^2 \quad (7)$$

Only the last term of (6) is kept as the first constants will cancel out.  $\Delta\beta_0 + \Delta\beta_1 S_t$  becomes 0 as the terms do not change with time.

$$L_{H0} = \frac{1}{2\sigma^2} \sum_{t=1}^T (\Delta R_t)^2 \quad (8)$$

This contrasts with the alternative where instability is implied. We assume  $\beta_t - \beta_0$  is approximated by the Gaussian random walk and  $\Delta R_t$  is therefore a Gaussian moving average of order 1 MA(1) with the specification:  $\Delta R_t \sim \eta_t + \psi_\eta \eta_{t-1}$ ,  $\eta_t \sim iidN(0, \sigma_\eta^2)$ , constant  $\psi_\eta < 1$ . Using the same *i.i.d* assumption we obtain the likelihood of this alternative process:

$$L_{HA} = \frac{1}{2\sigma^2} \sum_{t=1}^T \eta^2 \quad (9)$$

The qLL statistic is obtained by subtracting  $L_{HA}$  from  $L_{H0}$ :  $\frac{\sigma^2}{\sigma^2} \sum_{t=1}^T \eta^2 - \sum_{t=1}^T (\Delta R_t)^2$ . The test is therefore a variant of the Likelihood Ratio Test ( $LR_T$ ), so while it does not follow a chi-square distribution exactly, it does follow a certain related distribution that has its percentiles defined by Elliott and Müller (2006) and reported in their table, reproduced here as Table 4. The general extension to the  $LR_T$  can be made where we can reject the model related to the Null Hypothesis (the stable model) if the critical value is sufficiently negative. To obtain the  $\eta$  term it is necessary to follow additional matrix algebra and use the result of regressions (Elliott and Müller 2006). The appropriate steps are available in Appendix Section 9.2. While it is true that the underlying process can follow many different kinds of processes that could be very complex, we can detect their presence if we allow for their path to be approximated by a GRW (equivalent to having their first difference that follows an MA(1)).

## 4.2 CUSUM and CUSUM-squared tests

Separate tests to the one explained previously are considered. We explore the ones elaborated in Brown, Durbin, and Evans (1975) named the CUSUM and CUSUM-squared. These tests use the successive error terms of predictions of a standard Recursive-Least-Squares (RLS) model (details for this model in Section 9.2) that assumes stability of the  $\beta$  parameter (Young 2011). In this specific application, we use the result of the standard OLS as a baseline *prior* or initial value for the algorithm (ie.  $\beta_0$ ) and we use an empirical  $\hat{\sigma}_\varepsilon$  residual error that is based on the entire sample. By examining the prediction errors, one can observe whether or not they violate the  $N(0, \sigma^2)$  assumptions. Namely that they are a 1.) zero mean sequence  $E(\varepsilon_t) = 0$  and 2.) that they are serially uncorrelated  $E(\varepsilon_t \varepsilon_j) = 0$  when  $t \neq j$ . The first step consists in running a standard RLS algorithm and obtaining one-step-ahead errors from it. These errors are obtained as follows:

$$u_t = R_t - E(\hat{R}_t | R_{t-1}) \quad (10)$$

Effectively the realized price shock minus the predicted price shock at each observation. A transformation of these errors enables us to obtain a homoscedastic series.

$$u_{n,t} = \frac{u_t}{(1 + S_t^T P_t S_t)^{0.5}} \quad (11)$$

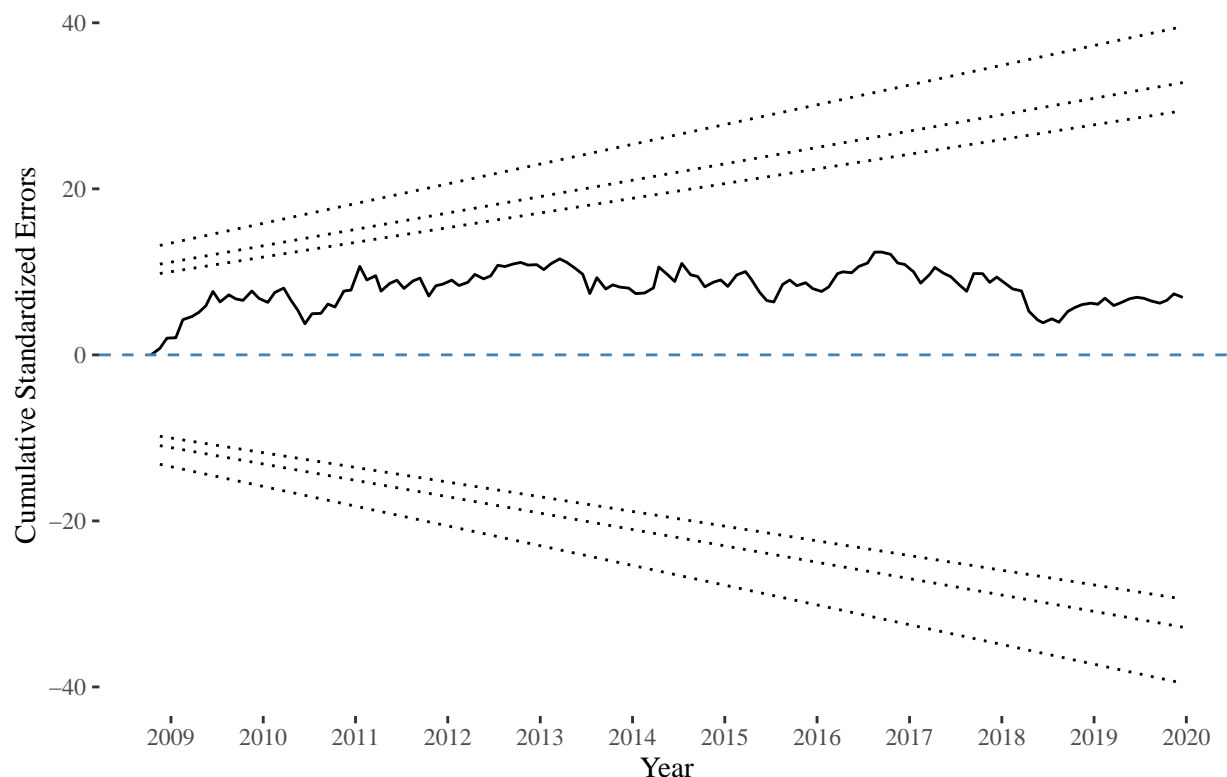


Where  $P_t$  is the covariance (or variance if there is only one news element at a time) of the  $\beta_t$ . A new series of summed up and standardized for use in the CUSUM test.

$$W_t = \frac{1}{\hat{\sigma}_{cs}} \sum_{i=k+1}^t u_{n,i} \quad (12)$$

In theory, these successively compounded errors should not stray too far from the zero-line if the true  $\beta_t$  is constant. We also use the confidence bands suggested by Brown, Durbin, and Evans (1975). They are constructed by constructing pairs of lines starting at time  $k$ :  $\pm a(T-k)^{0.5}$  and ending at time  $T$ :  $\pm 3a(T-k)^{0.5}$

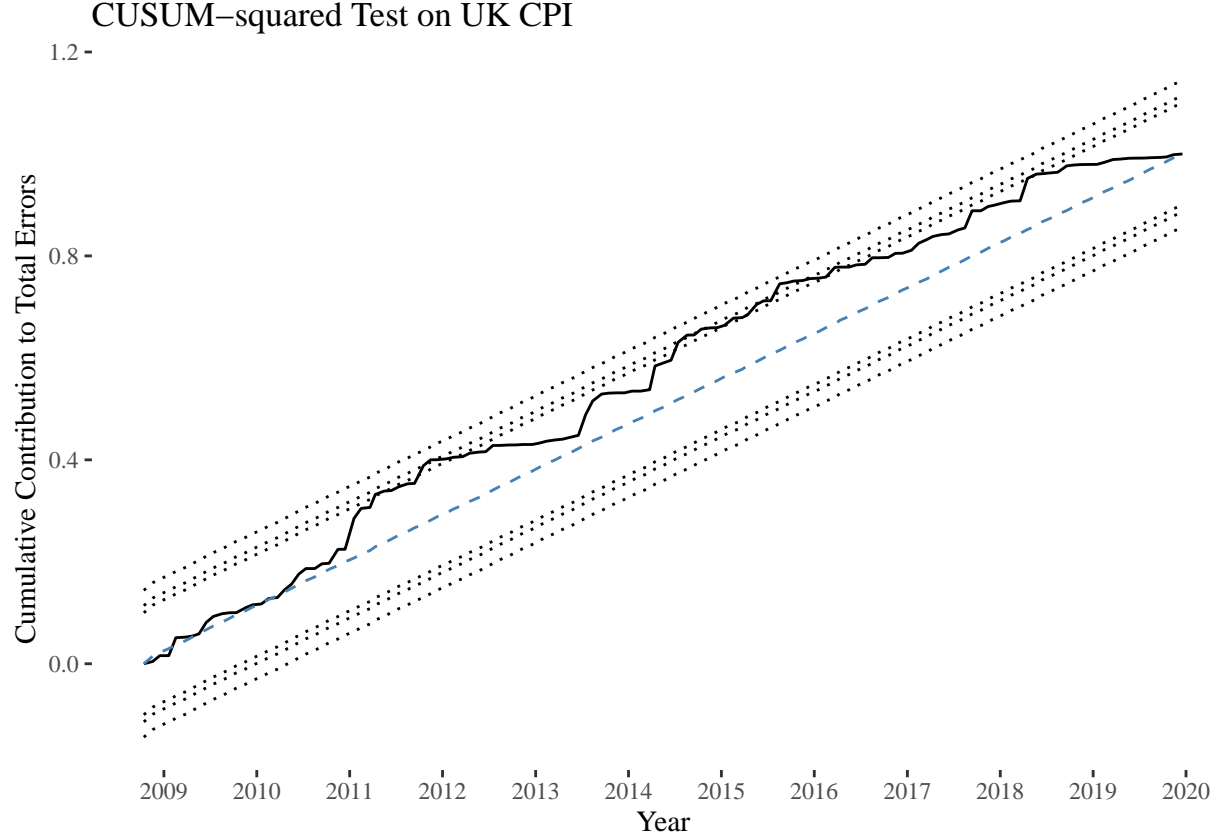
### CUSUM Test applied on UK CPI



The second test, the CUSUM-squared can be run by once more by creating a new series, similar to the  $W_t$  from earlier. Here we consider:

$$V_t = \frac{\sum_{i=k+1}^t u_{n,i}}{\sum_{i=k+1}^T u_{n,i}} \quad (13)$$

With a Null-Hypothesis of a constant parameter(s), these cumulative sum of squares should follow a beta distribution and its mean should be  $(k-h)/(N-h)$ . As significance levels, we use the bounds set as:  $\pm c_0 + (k-h)/(N-h)$ . The values of  $c_0$  depend on the sample size. Our sample sizes are typically larger than the maximum value given in Brown, Durbin, and Evans (1975). For those cases where



## 5 Parameter Path Estimations

Having established that there is instability over time in the market reactions to at least some of the news, we take on the task of obtaining a time series that conveys the change over the timeframe considered. Specifically, we would like to obtain a visual representation of the change of the parameter such that a researcher or investor is able to easily gather information from.

Two different algorithms from literature are utilized and both make different assumptions. Both are useful to help understand the possible sources of instability, to compare the different announcements and to make short-term forecasts. However, the main fundamental difference is that the WAR estimation uses the entire sample while the “simpler” STVP will only use outcomes that occurred in the past for its estimation.

### 5.1 Weighted Average Risk Minimization

Conveniently, this method is somewhat of a natural extended use of the qLL test elaborated in Section 4.1. It provides a heuristic means means to approximate a parameter path (Müller and Petalas 2010).

First, we consider our stable model which is analogous to the linear regression model that is presented in the Equation (2). The resulting  $\beta_{MLE/OLS}$  parameter obtained either through maximum likelihood estimation or minimization of ordinary least squares (depicted as the dot-dashed line in Figure 3) is representative of a scenario where it is assumed there is no time-related evolution of market reaction to news. As seen before, the likelihood in this case can be developed as Equation (6) or also written more generally as  $\sum_{t=1}^T l_t(\theta)$  with  $\theta$  containing the constant parameters.

We then turn to the case of varying  $\beta_t$ . Here, the general likelihood is the same but with time varying parameter(s) contained in  $\theta_t = \theta + \delta_t$  in  $t = 1, \dots, T$ . The  $\delta_t$  can be imagined as the vertical distances in Figure 3) between the constant  $\beta$  case and the *true* parameter path at time  $t$  that is unknown and that we are

attempting to approximate. The main argument for the method is that this general likelihood function for a time varying model can be approximated by a second-order Taylor expansion of the likelihood function around  $\beta_{OLS/MLE}$  and as a result an approximate estimate of the  $\delta_t$  term mentioned earlier is obtainable. Effectively it is demonstrated in the paper that the approximation of the log-likelihood function of the parameter path can be restructured such that a log-likelihood function of a gaussian random variable is recognizable and results in a “pseudo model” as:

$$\beta_{MLE} = \beta_t + T^{-1/2} \hat{H}^{-1} v_0 \quad (14)$$

$$s_t(\beta) = \hat{H} \delta_t + v_t, t = 1, \dots, T \quad (15)$$

$\hat{H}$  is the Hessian of the Taylor approximation divided by the sample size  $T$ ,  $s_t$  is the score function for the stable model. The exact derivations to obtain (15) are elaborated in Müller and Petalas (2010). The  $\delta_t$  vertical distances can be isolated in the latter expression as all the other elements are available to us.

To use the model of the authors and apply it to our case, we find expressions for  $s_t$  and  $\hat{H}$ . Essentially, the score function  $s_t$  and Hessian  $\hat{H}$  can be extracted from the stable linear regression shown in section Section 2 and the likelihood function is therefore the one seen in equation (7). Taking the derivative of the log likelihood once and then twice with respect to  $\beta_1$  yields the following two expressions respectively:

$$s_t = \frac{dl}{d\beta_1} = \sigma^{-2} (R_t - \hat{\beta}_0 - \hat{\beta}_1 S_t) S_t \quad (16)$$

$$\hat{H} = \frac{1}{T} \sum_{t=1}^T h_t(\hat{\beta}) = -\frac{dl^2}{d^2\beta_1^T} = \sigma^{-2} \sum_{t=1}^T S_t^2 \quad (17)$$

In the methodology, we are interested in the score of each particular observation and an *average* of the Hessians. This is why  $\hat{H}$  is used instead of  $h_t$  in (17). These two estimations, alongside the original  $\beta_1$  and  $\sigma^2$  from the time-invariant model are all that is needed to apply the entire WAR algorithm. The  $s_t$  series provide a direction that the parameter can take. However, the magnitude of these changes from one step to the next still needs to be explained.

Importantly, the score function is measuring the rate of change of the likelihood function with regards to its parameters. In the stable OLS model a lower valued score function for a certain observation in a sample indicates that a particular observation is decreasing the final  $\beta_{OLS}$  estimate. In the WAR minimization method illustrated by our UK CPI example in 3, the scores of the first few observations were mostly negative and led to  $\beta_t$  values below the  $\beta_{OLS}$ . Moreover, we recognize the Fisher Information of the linear regression case when taking the negative expectation of the Hessian (17) or equivalently taking the expectation of the square of (16). Intuitively, the  $s_t$  scores are therefore providing the vertical direction and a relative distance departure from the  $\beta_{OLS}$  whereas the  $\hat{H}$  Hessian is measuring the *overall* uncertainty

In Figure 3, one can observe 11 different random walk functions as well as the final estimated path for the UK CPI example. While they are all constructed using the scores and Hessian seen in (16) and (17), they each have a unique end point standard deviation. The weights in step (e) of the Appendix Section 9.3 depend on the sample size and the qLL statistic that are obtained using the qLL test statistic from Section 4.1. Specifically, it can be drawn from the expression  $\exp^{-\frac{1}{2} qLL(c_i)}$  that a higher standard deviation in a RW function leads to a smaller qLL statistic and therefore more importance is placed on the weight associated with that particular RW. However, this is balanced by the preceding expression  $\sqrt{T(1 - r_i^2)r_i^{T-1}/(1 - r_i^{2T})}$  that has the opposite effect of simply placing more weight on smaller RW functions. In our example, the random walk function that contributes the most weight for the UK CPI case is the 4th one.

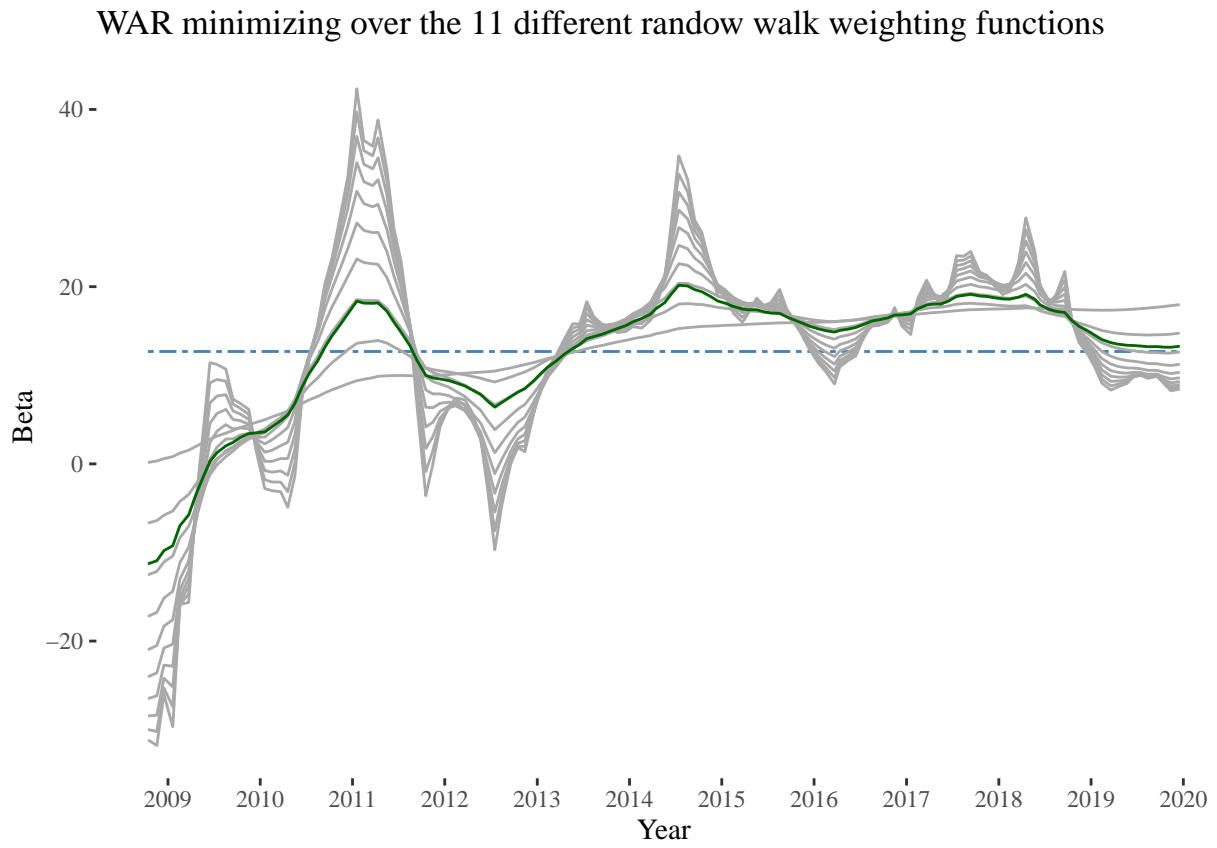


Figure 3: WAR minimizing over the 11 different random walk weighting functions

## 5.2 Standard Recursive Time Variable Parameter Algorithm (STVP)

Until now, the WAR minimization method provided a means to analyze the price reactions to the news over an entire sample period. In practice, it may be more useful to have a self-updating method that iteratively uses every new observation to improve the estimation progressively. The emphasis of the in this case is placed in gauging investor reactions to news in the present. When building this model, the assumption is that information available at time  $t$  includes everything before that moment in time and nothing more, specifically information set  $\Omega_{t-1}$ . This is in contrast with the WAR minimization and is why this particular method is also included.

Starting once more from the baseline case introduced in section Section 2, a time-dependent measure of market reaction to news  $\beta_t$  can be constructed by 1.) Transforming the problem of Ordinary Least Squares so that it can be recursively solved and 2.) Allowing for the parameter(s) to change over time sequentially.

To obtain the STVP algorithm presented in Section 9.3 the first step consists of replacing the OLS equations by their recursive counterparts and obtain the intermediate SRLS algorithm. The following equations are similar to the familiar OLS equations setup.

$$R_t = S_t^T \beta_1 + e_t \quad (18)$$

The error vector contains a series of random variables  $e_t$  that are assumed to be normally distributed with 0 mean and serially uncorrelated. We omit the constant intercept; the reasoning is elaborated in Section 7.

$$\hat{\beta}_{OLS} = (S^T S)^{-1} S^T R \quad (19)$$

$$\beta_t^{OLS} = \beta_{t-1}^{OLS} \forall t \quad (20)$$

When an observation  $y_j$  is added to the samples  $i = 1, \dots, N$ , an update to the  $\hat{\beta}$  is the same as re-estimating the entire sample with the new observation. Effectively, the “intermediate model” or Stochastic Recursive Least Squares Algorithm (SRLS) is thereafter obtained (the exact step-by-step is in Section 9.3). While this is a powerful tool, it does not necessarily help estimate  $y_j$  effectively. It is giving just as much weight to realizations of  $y$  early in the sample than those that immediately precede the new extra observation  $j$ . The contribution of every marginal observation is smaller and smaller as the sample grows. The  $\hat{\beta}$  converges to the full sample parameter estimator  $\hat{\beta}_{OLS}$  as seen in Figure 4.

Additional data points in the aforementioned methodology add information to the model and improve the accuracy of the  $\hat{\beta}$  estimator but it is always assumed that there is only one true  $\beta$  regardless of time  $t$ . In order to attenuate this assumption and include the flexibility that allows for a time dependence in a  $\beta_t$ , we introduce the random walk disturbance term mentioned in Section 2. The error term is now characterized by the incremental steps of a random walk. In bayesian terms, the previous estimation of  $\beta_t$  using the random walk becomes the *prior*.

$$R_t = S_t^T \beta_t + e_t \quad e_t \sim \mathcal{N}(0, \sigma^2) \quad (21)$$

$$\hat{\beta}_t = E(\beta_t | \hat{\beta}_{t-1}) \quad (22)$$

$$\beta_t = \beta_{t-1} + \eta_{t-1} \quad \eta_t \sim \mathcal{N}(0, \sigma_\beta^2) \quad (23)$$

The equations above are thus the foundation for the construction of the algorithm that is implemented for use and shown in Section 9.3. When initializing the first values, we set the variance of the estimator  $P_t^*$  to be very large and the  $\beta_t$  parameter to be 0 to reflect 0 knowledge of the nature of the estimation at the beginning of the sample.

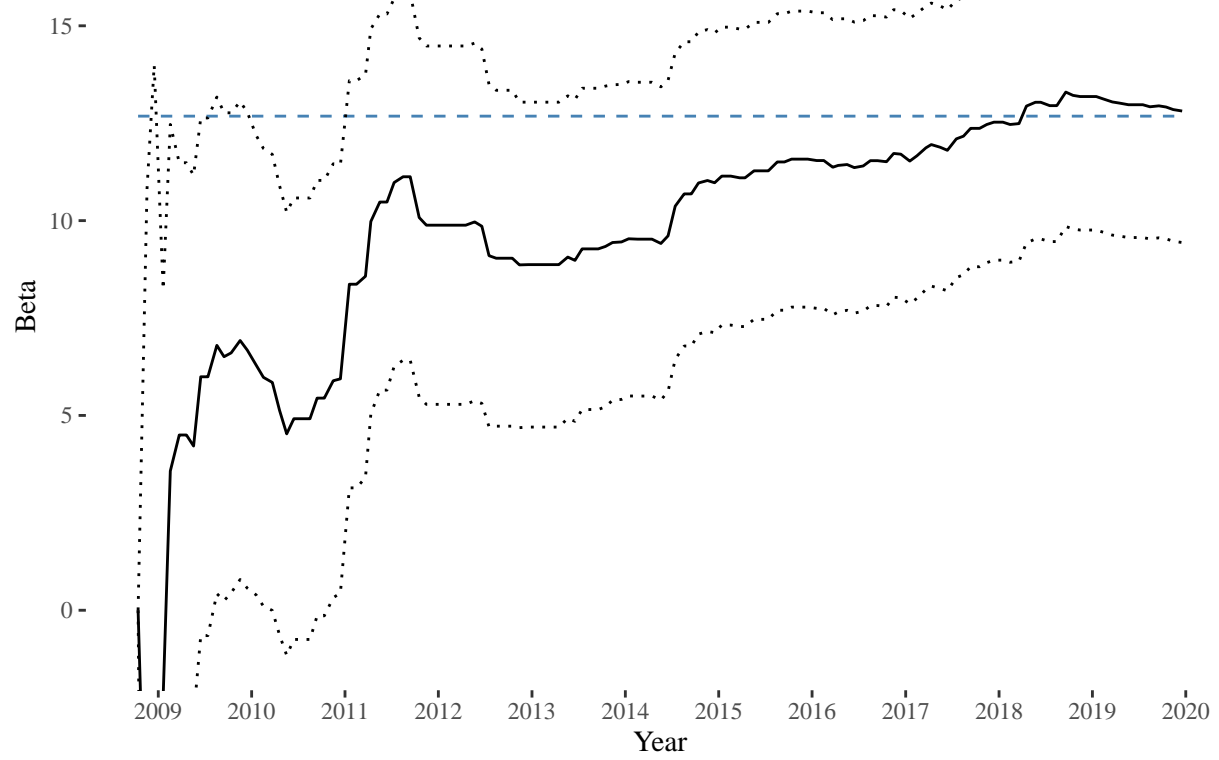


Figure 4: SRLS applied to the UK CPI

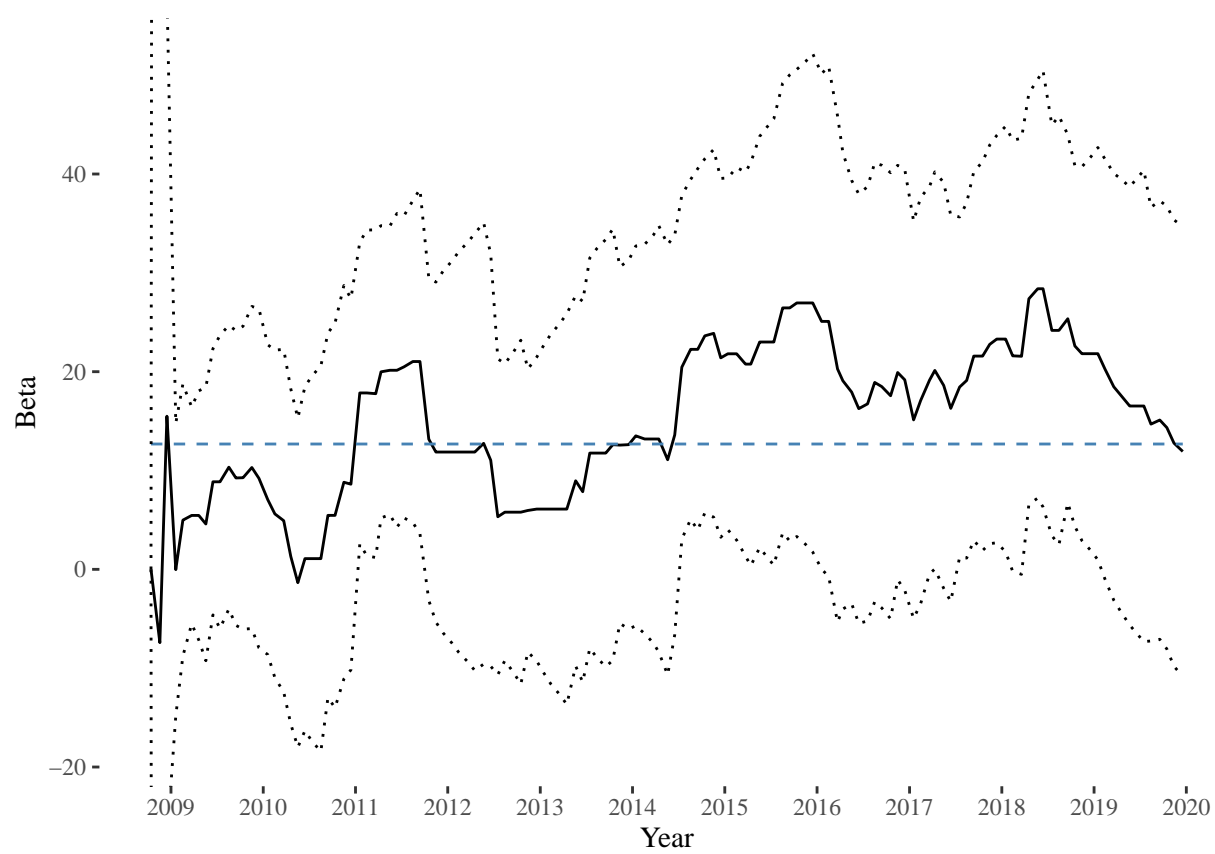


Figure 5: STVP applied to the UK CPI

## 6 Discussion

### 6.1 Differences and Irregularities of Two Paths

It is quite evident the STVP path is “slower to react” to the changes of its parameter. Informally, the reason for the lag is simply that the method does not benefit from retrospective analysis and is using only previous information to make an estimation. The high credible intervals towards the beginning of the sample are conveying the uncertainty.

The WAR Estimation method is especially useful when trying to analyze the evolution of the parameter over the entire sample period. Because of the “forward and backward passes” in the algorithm seen in steps (b) and (c) of section Section 9.4, the beginning of the sample period is estimated with as much precision as its ending. The addition of a reverse order filter means that the resulting estimate is “smoothed over” twice. This is a positive feature because there may be occasionally be “noisy” erroneous data points or exceptions due to outside factors that are “smoothed” out by the dictating trend and nearby observations both before and after their occurrence. Moreover, there is an inherent restriction of  $\sum_t^T \delta_t = 0$ , meaning that the positive  $\delta_t$  deviations from the average parameter value or  $\beta_{OLS}$  are exactly counterpoised by the negative ones. The path will always have the original, stable, time-invariant estimator as an average.

On the contrary, the STVP estimation is susceptible to exceptional or isolated changes to the price-reaction to the macro-announcements at any time during the sample. The  $\beta_t$  path can drastically change when both the surprise  $S_t$  shock and the difference  $R_t - S_t^T \beta_{t-1}$  are large in absolute terms. A higher noise variance  $\hat{\sigma}^2$  (constructed using residual errors) reduces the amount of variation the parameter will take while increasing the distance of the bounds which is undesirable but reflects the amount of uncertainty there is.

### 6.2 Accuracy of the Estimations

Using the bayesian perspective, the accuracy of the WAR line is measured using 95% posterior probability intervals (with priors being the weighting function of the random walks each having one of the 11 end-point variances.) (Müller and Petalas 2010). Therefore, given that the prior is correct and representative of the stable scenario case, the  $\beta_t$  lies within the interval zone with a 95% probability. The posterior distribution at each time  $t$ , is a weighted compromise of the different random walks and is the variance of this distribution that is used for the credible intervals.

Conversely, the STVP estimation at time  $t$  is achieved using the previous observation at  $t - 1$  as a prior. One can observe that the STVP estimation bounds are typically much more narrow than that of the WAR.

### 6.3 One-Step ahead Forecasts

While the primary aim of the study was not to obtain forecasts, we investigate the possibility of a naive prediction with their associated prediction intervals. The term *naive* is used because it will be assumed that the best measures for expectation and variance for the 5 minute pip change of  $\hat{R}_{t+1}$  are obtained by using the most recent estimates of the sample at time  $t$ . Furthermore, for this forecast, we put ourselves in the situation of an investor who has just observed the release of the UK CPI on January 15th 2020 at 11:30:00, and would like a measurement of certainty of where price will lie less than 5 minutes afterwards. Information set  $\mathcal{F}_t$  includes all the past events whereas  $S_{t+1}$  is the new news surprise figure.

$$E(\hat{R}_{t+1}|\mathcal{F}_t, S_{t+1}) = \hat{\beta}_t S_{t+1} \quad (24)$$

$$Var(\hat{R}_{t+1}|\mathcal{F}_t, S_{t+1}) = S_{t+1}^2 \sigma_{\beta}^2 + \sigma_{\epsilon}^2 \quad (25)$$

Using the variances for the  $\beta_t$  parameters of each model allows for creation of the two different 95% prediction intervals. These intervals are much wider than the credible intervals considered earlier as there is the additional uncertainty of the measurement  $\sigma_{\epsilon}^2$  or often referred to as the irreducible error. This error is quite



large relative to the  $\sigma_\beta^2$ , a clear sign that the models could be improved possibly by considering additional predictor variables aside from  $S_t$ .

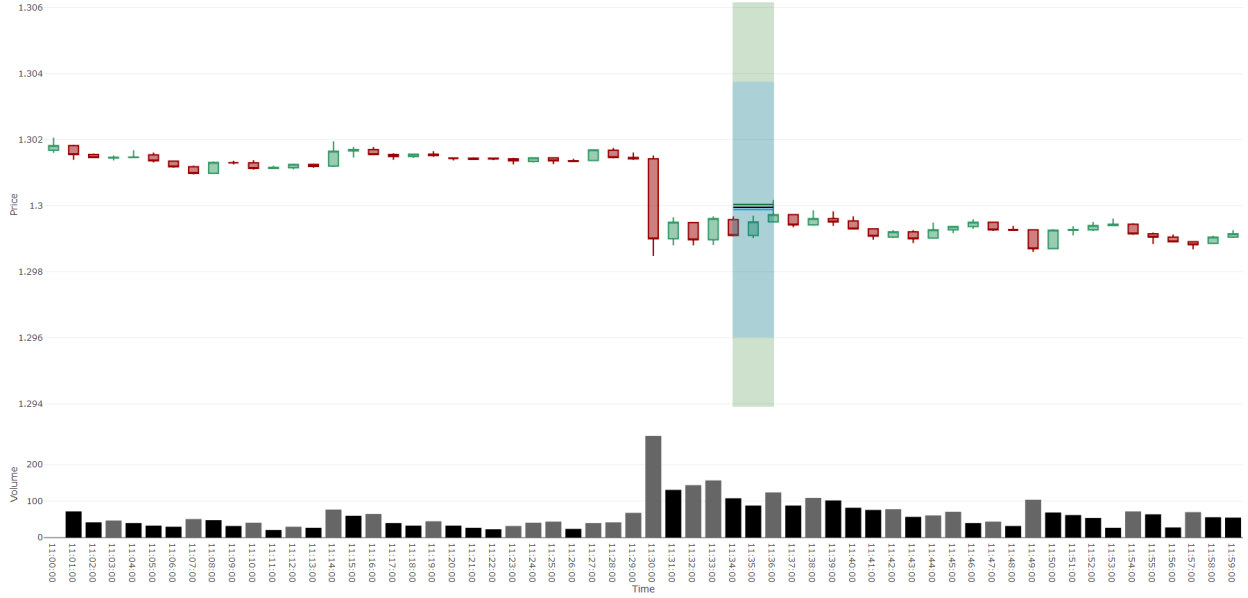


Figure 6: GBP/USD on the 15th of January between 14h00 and 15h00 GMT+2. UK CPI released with a figure of 1.3% versus the expected 1.5% (equivalently a -1.1579697 Std. Dev shock). Point-Estimates and 95% prediction bounds are colored in blue and green for the WAR and STVP estimations respectively. The time-invariant OLS case is also added in black for reference.

## 6.4 Use of the Post-Announcement 5 Minute Window

The erroneous or exceptional data can sometimes be quite obvious. An example would be that of the Canadian Core Retail Sales in July of 2013. On this day, the figure achieved a 1.2% increase versus the 0.1% that was expected and yet the 5 minute post-price difference was merely -11.3 pips. The price move was instantaneous to the extent that the beginning of the 14:30:00 candle already marked the majority of the price shock. A solution to this issue would have been to consider using a price before 14:30:00 although this would have introduced the issue of finding a general timing for all other releases that does not include irrelevant price movements before to the announcement time-stamp. Furthermore, in some circumstances, it could be desirable to omit these price jumps as they would result in what is called “slippage” (A difference between the price an investor will place an order at and the price it is executed at).

There are several explanations for these exceptions. One example in particular would be market interventions of the central banks which have been proven to occur during or in time proximity to figure releases as shown in studies such as that of Dominguez (2003). Similarly, there is evidence of insider trading preceding news announcements. Assets traded in such as the E-mini S&P 500 futures where exact order imbalance can be measured. It is perfectly possible that the same kind of activity is taking place in the decentralized currency exchange (or FX) markets (Bernile, Hu, and Tang 2016). It is impossible (or extremely difficult) to detect these external reasons for price shocks and the WAR conveniently “smooths” them over.

One can observe that the estimation can “stray away from the dictating trend” to a larger extent than the other method.

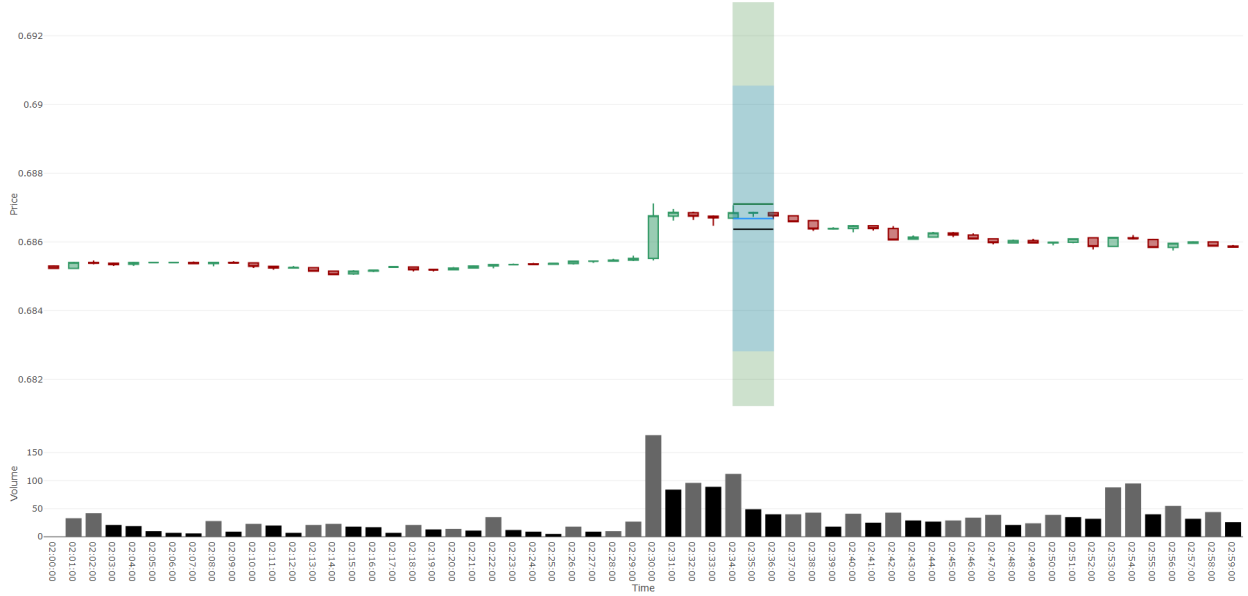


Figure 7: AUD/USD on the 10th of January between 02h00 and 03h00 GMT+2. Retail Sales announced at 0.9% versus the expected 0.4% that was expected (equivalently a  $+0.8801813$  Std. Dev shock). Point-Estimates and 95% prediction bounds are colored in blue and green for the WAR and STVP estimations respectively. The time-invariant OLS estimate is also added in black for reference.

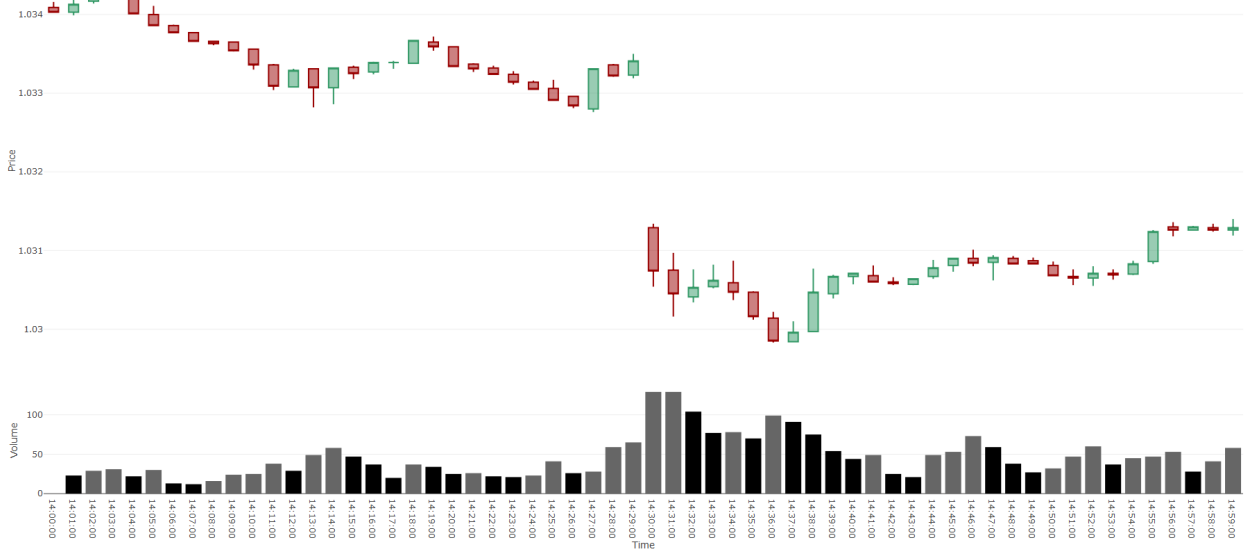


Figure 8: USD/CAD on the 27th of July between 14h00 and 15h00 GMT+2. The “missing” data can be attributed to the sudden change in price and lack of liquidity in the moments preceding 14:30:00.

## 7 Appendix

### 7.1 Parameter Paths

The following path estimations should be interpreted as follows: a one standard deviation unanticipated positive change in the economic variable  $S_t$  results in an appreciation or depreciation in the exchange rate of the relevant currency pair by  $R_t$  amount of pips or basis points.

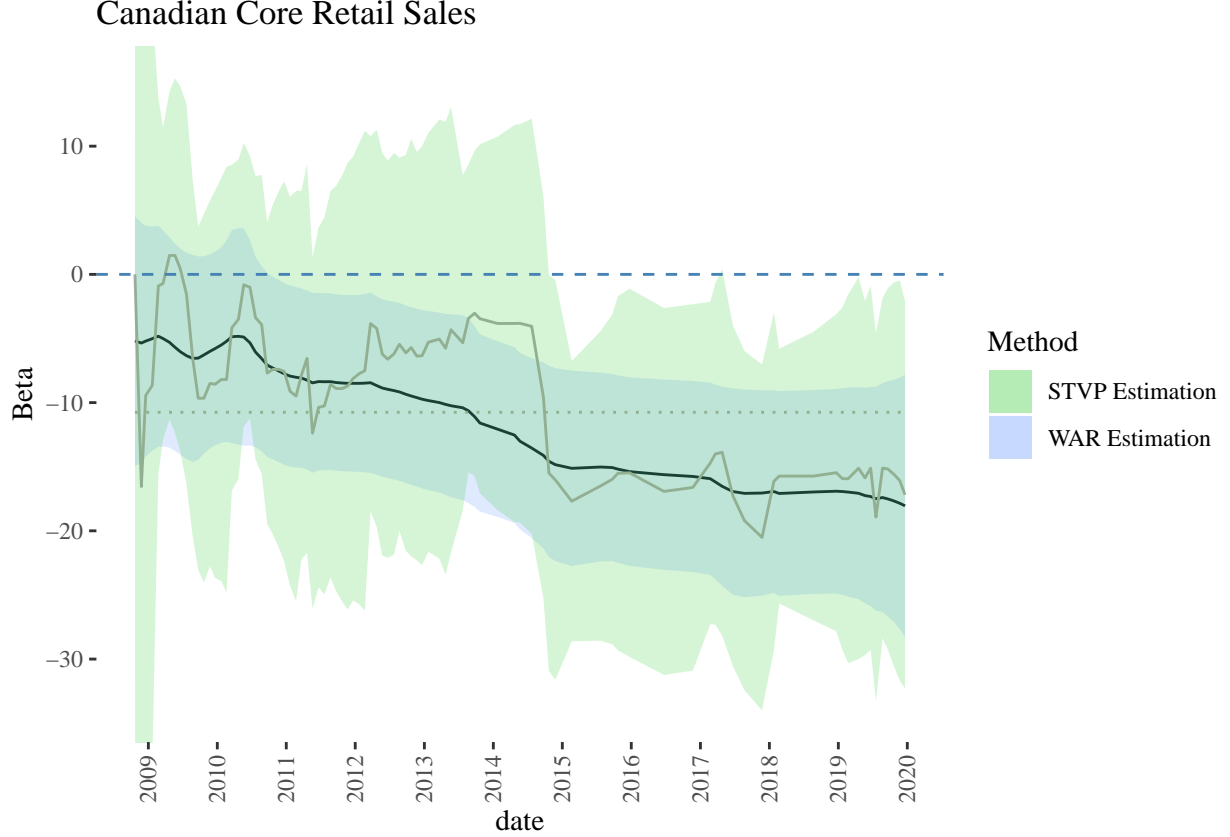


Figure 9: Note: 1 Std Dev is +0.605% greater than expected. Currency pair: USD/CAD.

### 7.2 Stochastic Recursive Least Squares Algorithm (SRLS)

1.  $\hat{\alpha}_t = \hat{\alpha}_{t-1} + g_t(R_t - S_t^T \hat{\alpha}_{t-1})$
2.  $g_t = P_{t-1}^* S_t (\hat{\sigma}^2 + S_t^T P_{t-1}^* S_t)^{-1}$
3.  $P_t^* = P_{t-1}^* - g_t S_t^T P_{t-1}^*$

### 7.3 Standard Recursive Time Variable Parameter Algorithm (STVP)

In the case of this study, it is assumed that the parameter follows a simple random walk. As a result,  $A = D = I$ . A prior value that represents a 25% change per observation is chosen for the diagonals of  $Q_a$ .

Prediction (Prior)

1.  $\hat{\alpha}_t | \hat{\alpha}_{t-1} = A \alpha_{t-1}$
2.  $P_t^* | P_{t-1}^* = A P_{t-1}^* A^T + D Q_a D^T$

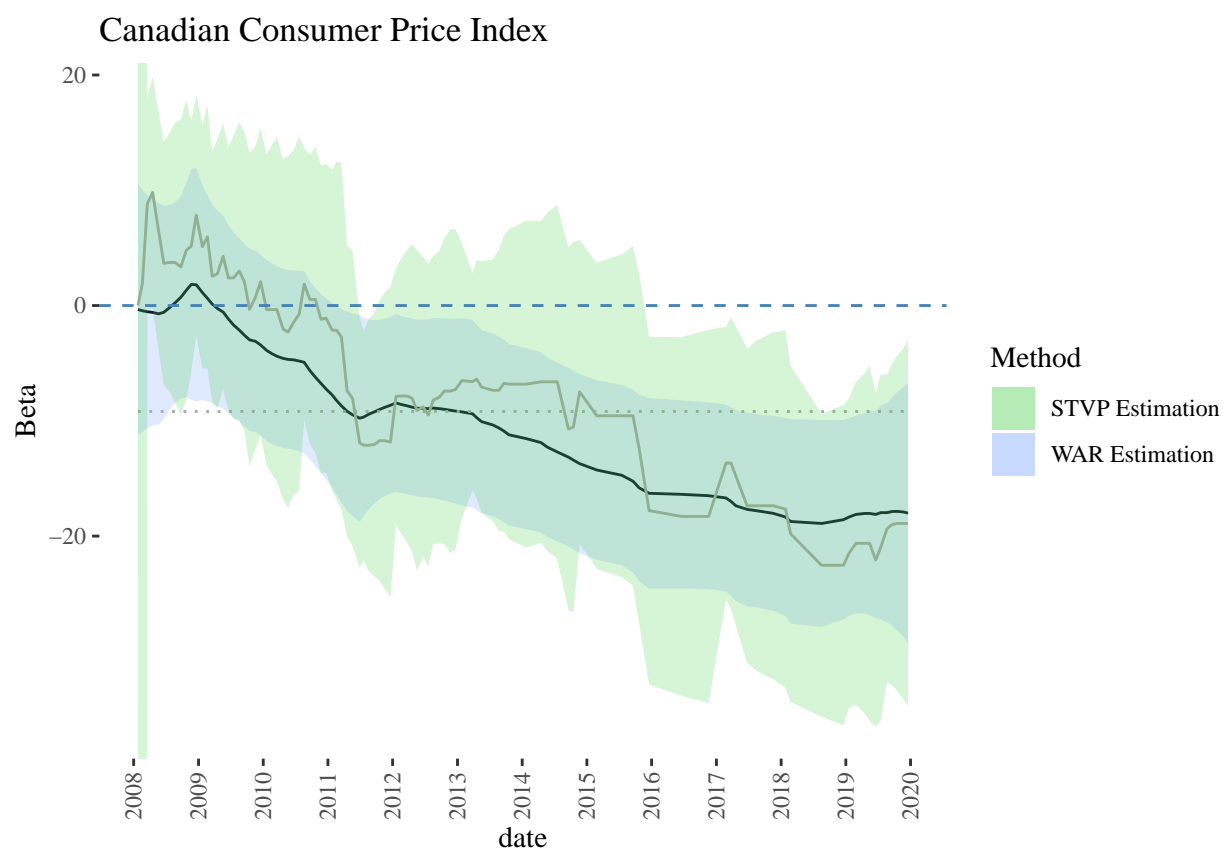


Figure 10: Note: 1 Std Dev is +0.229% greater than expected. Currency pair: USD/CAD.

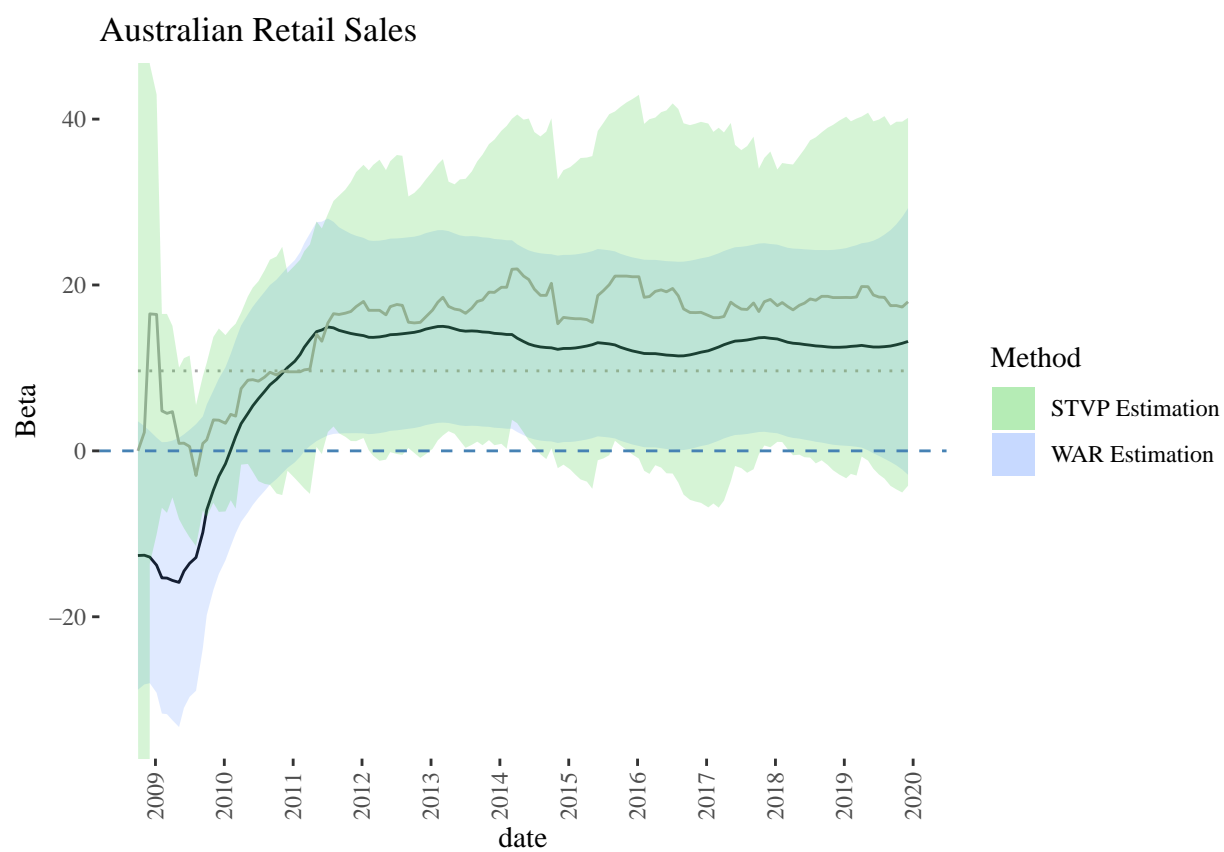


Figure 11: Note: 1 Std Dev is +0.569% greater than expected. Currency pair: AUD/USD.

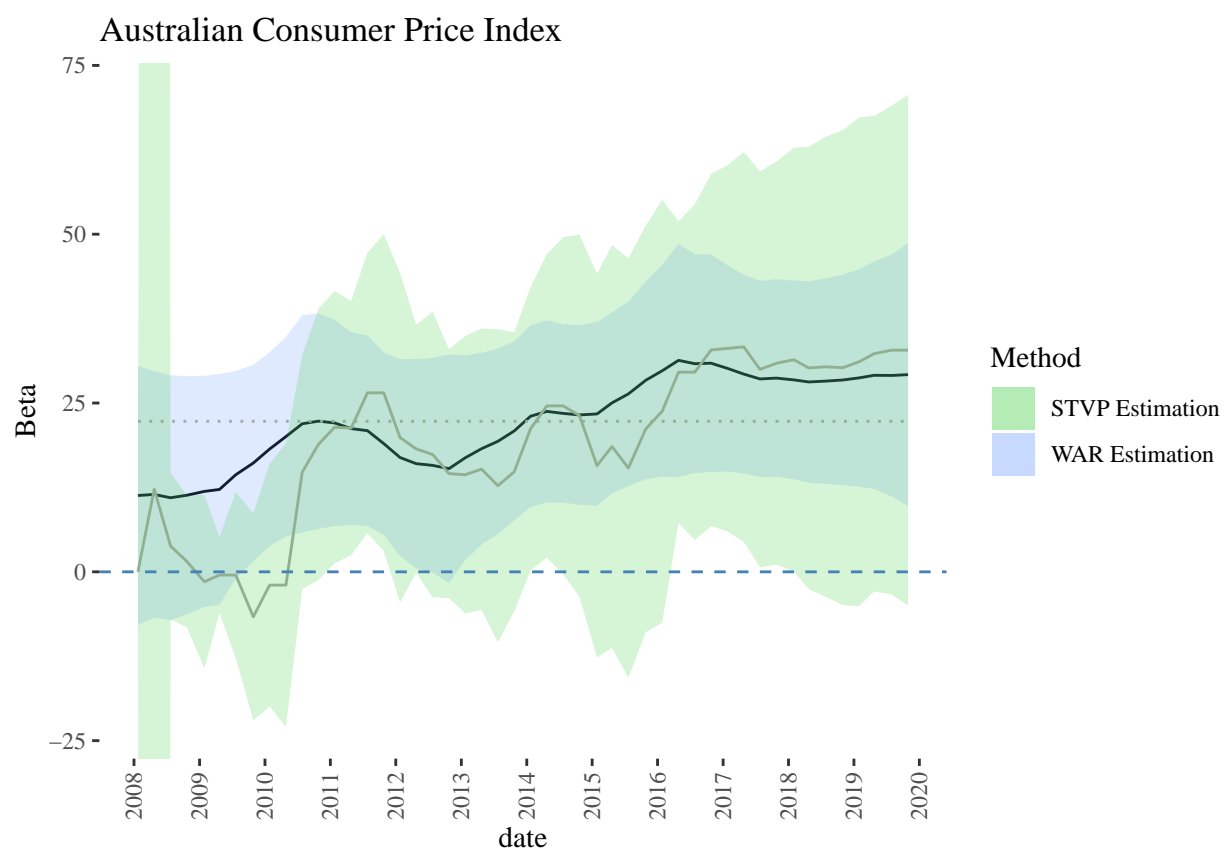


Figure 12: Note: 1 Std Dev is +0.229% greater than expected. Currency pair: AUD/USD.

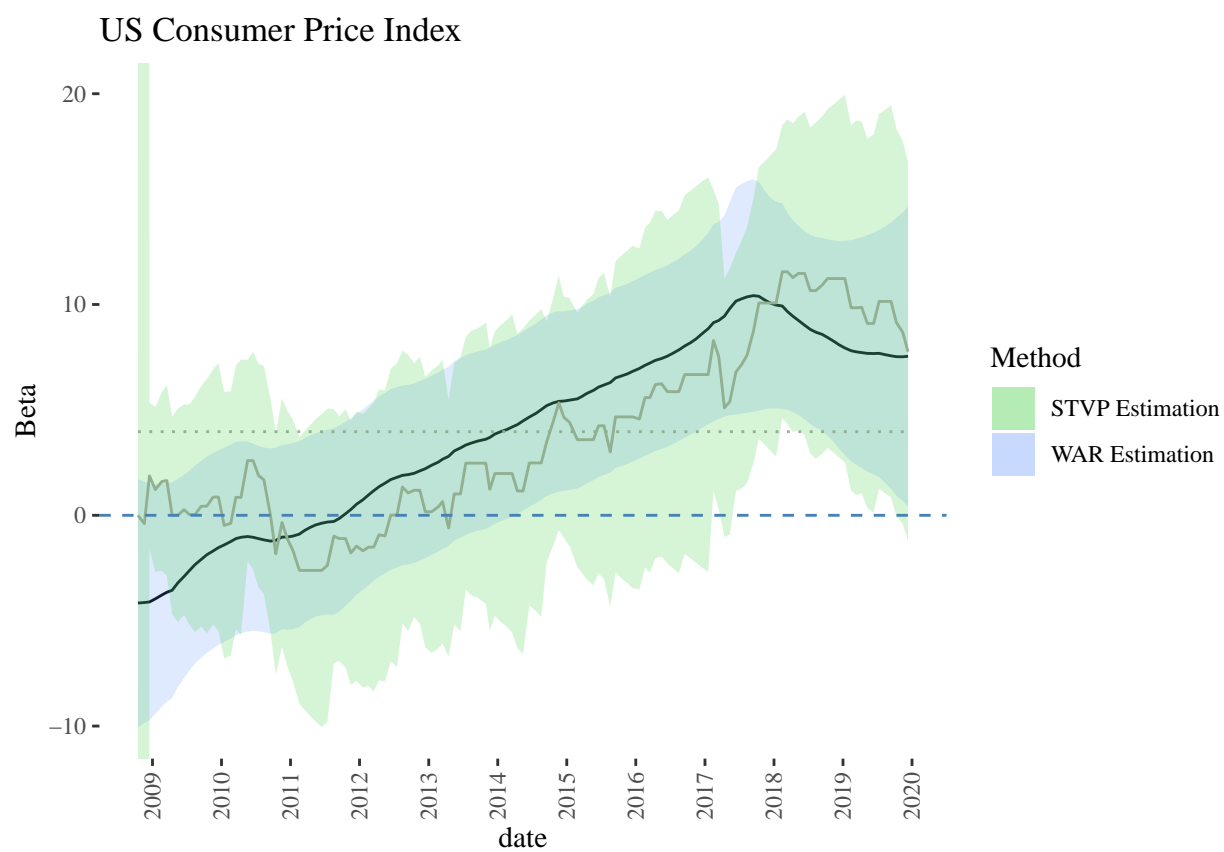


Figure 13: Note: 1 Std Dev is +0.117% greater than expected. Currency pair: USD/CHF.

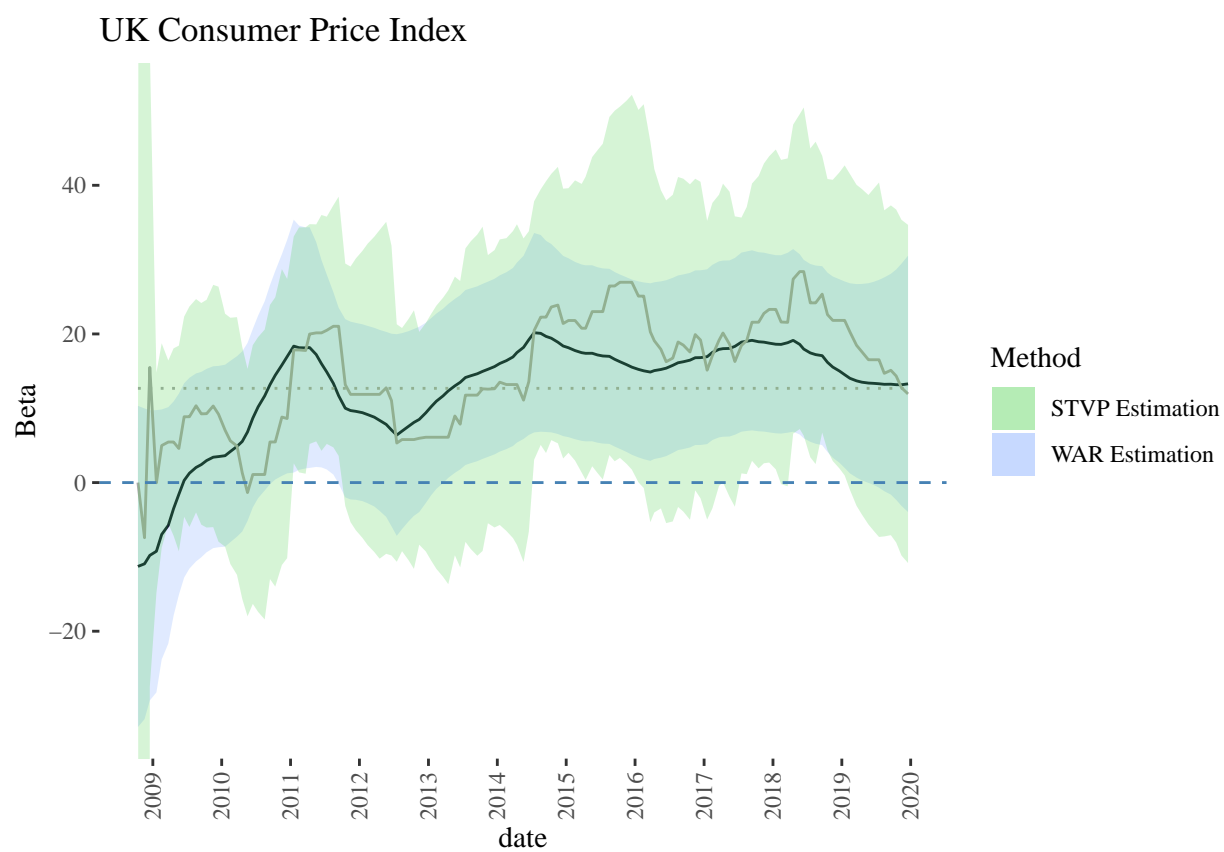


Figure 14: Note: 1 Std Dev is +0.173% greater than expected. Currency pair: GBP/USD.



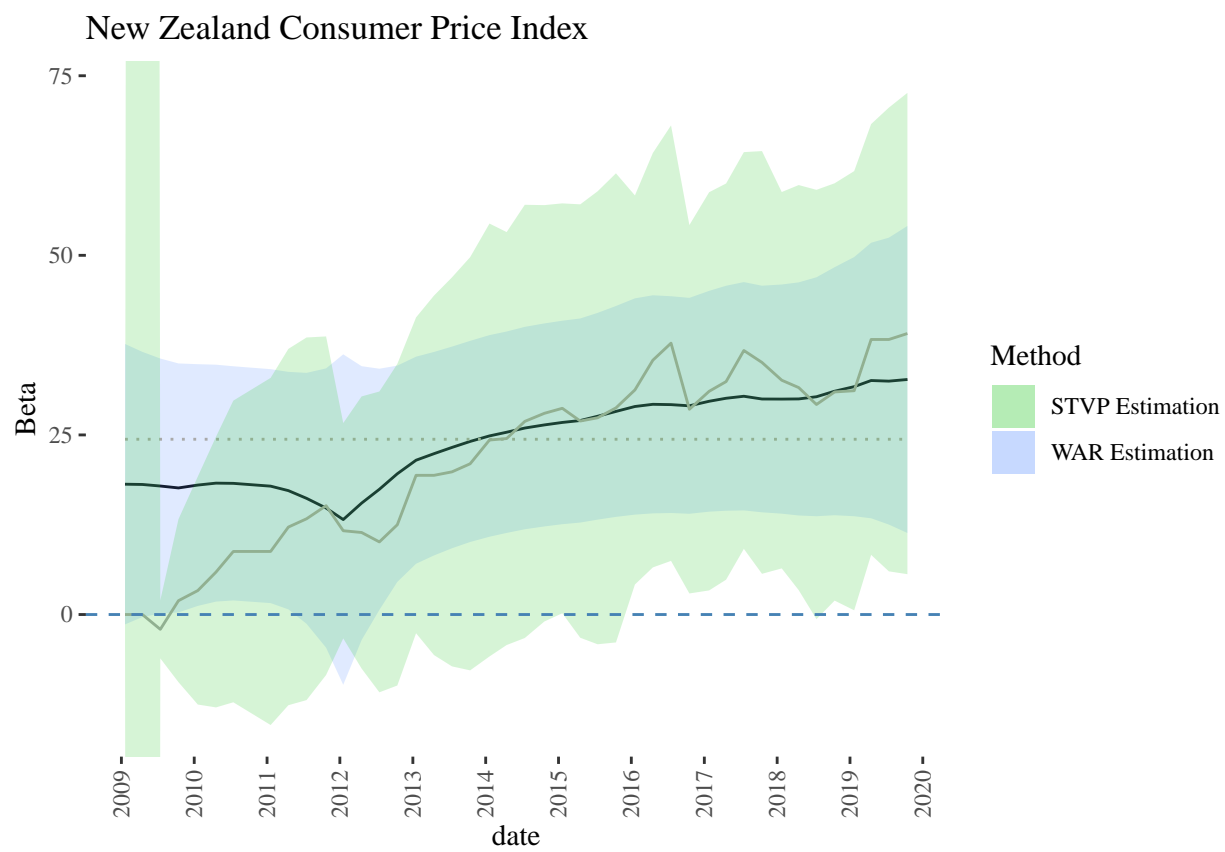


Figure 15: Note: 1 Std Dev is +0.206% greater than expected. Currency pair: NZD/USD.

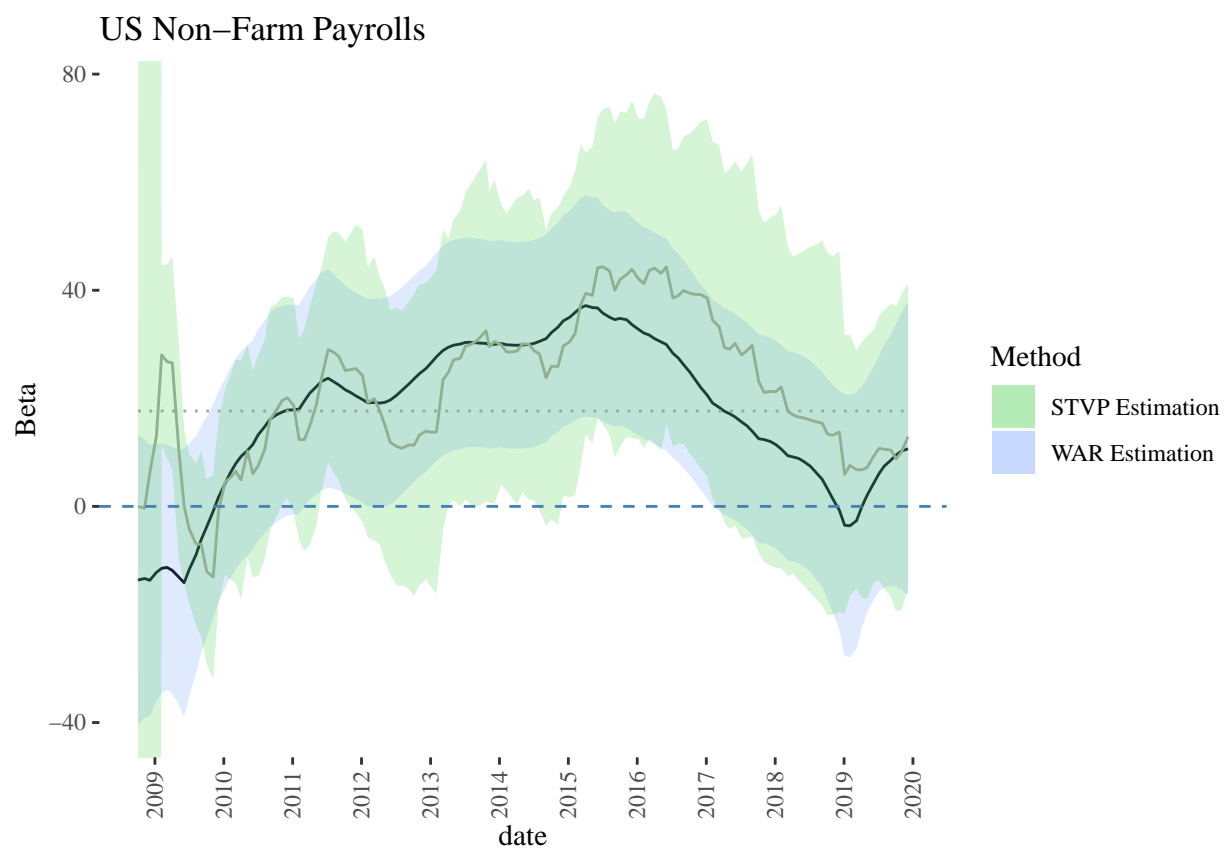


Figure 16: Note: 1 Std Dev is +66.5 thousand more people than expected. Currency pair: USDCHF.

Correction (Posterior)

3.  $\hat{\alpha}_t = \hat{\alpha}_{t-1} + g_t(R_t - S_t^T(\hat{\alpha}_t|\hat{\alpha}_{t-1}))$
4.  $g_t = (P_t^*|P_{t-1}^*)S_t(\hat{\sigma}^2 + S_t^T(P_t^*|P_{t-1}^*))^{-1}$
5.  $P_t^* = P_t^*|P_{t-1}^* - g_t S_t^T(P_t^*|P_{t-1}^*)$

## 7.4 Steps to obtain qLL statistic

1. Compute the OLS residuals  $\hat{\varepsilon}_t$  by regressing  $R_t$  on  $S_t, Z_t$ ;
2. Construct a consistent estimator  $\hat{V}_X$  of the  $k * k$  long-run covariance matrix of  $S_t \varepsilon_t$ . When  $\varepsilon_t$  can be assumed uncorrelated, a natural choice is the heteroscedasticity robust estimator  $\hat{V}_X = T^{-1} \sum_{t=1}^T X_t X_t' \varepsilon_t^2$
3. Compute  $\hat{U}_t = \hat{V}_X^{-1/2} X_t \hat{\varepsilon}_t$  and denote the  $k$  elements of  $\hat{U}_t$  by  $\hat{U}_{t,i}$ ,  $i = 1, \dots, k$ .
4. For each series  $\hat{U}_{t,i}$ , compute a new series,  $\hat{w}_{t,i}$  via  $w_{t,i} = \bar{r} \hat{w}_{t-1,i} + \Delta \hat{U}_{t,i}$ , and  $\hat{w}_{1,i} = \hat{U}_{1,i}$ , where  $\bar{r} = 1 - 10/T$ .
5. Compute the squared residuals from OLS regressions of  $\hat{w}_{t,i}$  on  $\bar{r}^t$  individually, and sum all of those over  $i = 1, \dots, k$ .
6. Multiply this sum of sum of squared residuals by  $\bar{r}$ , and subtract  $\sum_{i=1}^k \sum_{t=1}^T (\hat{U}_{t,i})^2$

## 7.5 Steps to obtain WAR minimization path

1. For  $t = 1, \dots, T$ , let  $a_t$  and  $b_t$  be the first  $p$  elements of  $\hat{H}^{-1} s_t(\hat{\theta})$  and  $\hat{H} \hat{V}^{-1} s_t(\hat{\theta})$  respectively.
2. For  $c_i \in C = 0, 5, 10, \dots, 50, i = 1, \dots, 11$  compute
  - (a)  $r_i = 1 - c_i/T$ ,  $z_{i,1} = x_1$  and  $z_{i,t} = r_i z_{i,t-1} + x_t - x_{t-1}, t = 2, \dots, T$ ;
  - (b) the residuals  $\{\tilde{z}_{i,t}\}_{t=1}^T$  of a linear regression of  $\{z_{i,t}\}_{t=1}^T$  on  $\{r_i^{t-1} I_p\}_{t=1}^T$
  - (c)  $\bar{z}_{i,T} = \tilde{z}_{i,T}$ , and  $\bar{z}_{i,t} = r_i \bar{z}_{i,t+1} + \tilde{z}_{i,t} - \tilde{z}_{i,t+1}, t = 1, \dots, T-1$ ;
  - (d)  $\{\hat{\beta}_{i,t}\}_{t=1}^T = \{\hat{\theta} + a_t - r_i \bar{z}_{i,t}\}_{t=1}^T$ ;
  - (e)  $qLL(c_i) = \sum_{t=1}^T (r_i) \bar{z}_{i,t} - a_t) \tilde{b}_t$  and  $\tilde{w}_i = \sqrt{T(1 - r_i^2) r_i^{T-1} / (1 - r_i^{2T})} e^{-\frac{1}{2} qLL(c_i)}$  (set  $\tilde{w}_0 = 1$ )
3. Compute  $w_i = \tilde{w}_i / \sum_{j=1}^{11} \tilde{w}_j$ .
4. The parameter path estimator is given by  $\{\hat{\beta}_t\}_{t=1}^T = \{\sum_{i=1}^{11} w_i \hat{\beta}_{i,t}\}_{t=1}^T$ .
5. The statistic qLL(10) tests the null hypothesis of stability of  $\beta$  and rejects for small values. Critical values depend on  $p$  and are tabulated in Table 1 of Elliott and Müller (2006) and 4.

## 7.6 Variance for credible intervals for the WAR path

$$\Omega_t = \sum_{i=1}^{10} w_i (T^{-1} \hat{S}_\beta \kappa_t(c_i) + (\hat{\beta}_{i,t} - \hat{\beta}_t)(\hat{\beta}_{i,t} - \hat{\beta}_t)'), \kappa_t(c) = \frac{c(1+e^{2c}+e^{2ct/T}+e^2 c(1-t/T))}{2e^{2c}-2}$$

## References

Andersen, Torben G, Tim Bollerslev, Francis X Diebold, and Clara Vega. 2003. "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange." *American Economic Review* 93 (1): 38–62.

doi:10.1257/000282803321455151.

Ben Omrane, Walid, Tanseli Savaser, Robert Welch, and Xinyao Zhou. 2019. “Time-Varying Effects of Macroeconomic News on Euro-Dollar Returns.” *The North American Journal of Economics and Finance* 50 (November): 101001. doi:10.1016/j.najef.2019.101001.

Bernile, Gennaro, Jianfeng Hu, and Yuehua Tang. 2016. “Can Information Be Locked up? Informed Trading Ahead of Macro-News Announcements.” *Journal of Financial Economics* 121 (3): 496–520. doi:10.1016/j.jfineco.2015.09.012.

Brown, R. L., J. Durbin, and J. M. Evans. 1975. “Techniques for Testing the Constancy of Regression Relationships over Time.” *Journal of the Royal Statistical Society: Series B (Methodological)* 37 (2): 149–63. doi:10.1111/j.2517-6161.1975.tb01532.x.

Dominguez, Kathryn M.E. 2003. “The Market Microstructure of Central Bank Intervention.” *Journal of International Economics* 59 (1): 25–45. doi:10.1016/S0022-1996(02)00091-0.

Elliott, Graham, and Ulrich K. Müller. 2006. “Efficient Tests for General Persistent Time Variation in Regression Coefficients.” *Review of Economic Studies* 73 (4): 907–40. doi:10.1111/j.1467-937X.2006.00402.x.

Goldberg, Linda S., and Christian Grisse. 2013. “Time Variation in Asset Price Responses to Macro Announcements.” *SSRN Electronic Journal*. doi:10.2139/ssrn.2321203.

Müller, Ulrich K., and Philippe-Emmanuel Petalas. 2010. “Efficient Estimation of the Parameter Path in Unstable Time Series Models.” *Review of Economic Studies* 77 (4): 1508–39. doi:10.1111/j.1467-937X.2010.00603.x.

Newey, Whitney K., and Kenneth D. West. 1987. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55 (3): 703. doi:10.2307/1913610.

White, Halbert. 2001. *Asymptotic Theory for Econometricians*. Rev. ed. San Diego: Academic Press.

Young, Peter C. 2011. *Recursive Estimation and Time-Series Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-21981-8.