WAGES AND PRICES OF FOREIGN GOODS IN THE INFLATIONARY PROCESS IN ICELAND

Ásgeir Danielsson
Wages and prices of foreign goods in the inflationary process in Iceland

Ásgeir Danielsson*
Central Bank of Iceland
September 2021

Abstract

In this paper we discuss the relationships between the CPI in Iceland, the unit labour cost, and the price of foreign goods, and their role in equations for forecasting inflation. We find that the logs of these variables are cointegrated, the cointegrating vectors are stable over many different data periods, and the coefficients satisfy the homogeneity condition. On the other hand, the coefficients in regressions of log difference of the CPI on log differences of the other variables, and a constant, are unstable, and for data for the last two decades, the homogeneity condition is always rejected. The coefficient for changes in unit labour cost, the price of the most important cost item, is often insignificant, while the constant, which shouldn’t be in the equation, is frequently highly significant. It is shown that the estimates of coefficients in the equation in log differences of the variables depend on the coefficients of correlations between the variables, and their standard deviations, which have diverged very much since the turn of the century. Large standard deviations of changes in unit labour cost, and especially of changes in the price of foreign goods, compared to standard deviations of changes in the CPI, contribute to lower coefficient estimates, and to the significance of the constant. In the paper we discuss how the long-run, cointegrated, relationship between the logs of the variables can be used to obtain valuable information for forecasting the rate of inflation. We also present estimation of an equation for the log difference in CPI where the error-correction term is the estimated error of an AR-equation for the errors from the equation in logs.

JEL: C130, E31, E44, E52, E65

Keywords: inflation, wage rates, cointegration, error-correction model, homogeneity condition, variation in variables

I thank Guðmundur Guðmundsson, Lúðvík Elíasson, and two anononymous referees, for valuable comments.

Author’s emails: ad@cb.is, asgeir.danielsson@gmail.com.
1. Introduction

Many people in Iceland, economists, business leaders, entrepreneurs, politicians, and others, have expressed their belief in a close relationship between increases in wages and the rate of inflation. On January 28, 2021, the Minister of Finance and Economic Affairs responded to an MP, who had asked in Parliament why Iceland was the only country in Europe where inflation increased during the Covid-19 crises, that he believed that this was the result of larger increases in wages in Iceland.\footnote{Reported on the news website Miðjan, https://www.midjan.is/bjarni-launahaeckkanir-valda-verdbolgu/} And in an interview published on May 20, 2021, the governor of the Central Bank of Iceland stated: "There is a causal relationship between wage rates and inflation ..."\footnote{https://www.frettabladid.is/markadurinn/sedlabankastjori-segir-ad-rikid-verdi-ad-fara-dragna-sig-i-hle/} Regular polling of firms show that the respondents think that the main reason for changing the firms’ prices are changes in costs, where wage cost is usually the largest component. (See e.g. CBI, 2019)

The wage cost is a very large share of total cost of most firms. Statistics Iceland estimates the aggregate annual wage cost in the period considered in this paper to be around 60% of gross factor income and around 50% of GDP. Given these facts, it is somewhat surprising that statistical analysis, using data for this century, does not detect a significant relationship between changes in wage rates and the rate of inflation. Estimation with OLS of a simple equation using data for the period 2003Q1-2017Q4 – a reasonable period given that inflation targeting framework for the monetary policy was introduced in 2001, and this is also the data period used for the estimation of the equation for inflation in Central Bank of Iceland’s (CBI’s) main macro model, QMM (see Danielsson et. al., 2019) – gives the following results (t-values in parenthesis):

$$Dlog(p_t) = 0.0104 - 0.0245 Dlog(w_t) + 0.0959 Dlog(pf_t) + 0.0342 D082$$ \hspace{1cm} (1)

Here $p_t$ is the CPI, $w_t$ is the trend unit labour cost (ULCT in QMM), and $pf_t$ is the weighted CPI in the main trading countries in ISK, Icelandic krónur (= EER · WPCI, where EER is the nominal price of foreign currencies, and WPCI is the weighted CPI in Iceland’s main trading partners). All indices are scaled to the average of unity in 2005. D082 is a dummy taking the value of unity in the second quarter of 2008, on the eve of the financial crisis. This simple equation has an $R^2$ of 0.59, and it passes standard tests for normality, homoskedasticity, and no autocorrelation of the residuals.

The estimate of the coefficient for changes in the price of foreign goods and services is highly significant. The constant, and the coefficient for the dummy are also highly significant, but the coefficient of changes in unit labour cost is slightly negative and insignificant. Adding lagged log differences in $w_t$ and $pf_t$ as explanatory variables increases $R^2$ to 0.76. The equation, which passes standard tests for no autocorrelation, homoskedasticity and normality of residuals, is:

$$Dlog(p_t) = 0.0098 - 0.0204 Dlog(w_t) + 0.0074 Dlog(w_{t-1})$$ \hspace{1cm} (8.49) \hspace{1cm} (0.120) \hspace{1cm} (1')

$$+ 0.0931 Dlog(pf_t) + 0.0772 Dlog(pf_{t-1}) + 0.0257 D082$$ \hspace{1cm} (6.17) \hspace{1cm} (4.10)

The coefficient for lagged log difference of prices of foreign goods is highly significant, as is the coefficient for contemporary log difference of prices of foreign goods, but the coefficient for contemporary log difference of unit labour cost, and that of lagged log difference of unit labour cost are far from being significant, and their sum is negative, -0.013.
The standard interpretation of these results is that changes in wage rates do not affect the rate of inflation, which seems somewhat unreasonable. The standard interpretation of the estimated coefficients of changes in prices of foreign goods, which is 0.170, is that the long-run pass through of these costs into consumer prices was 17%. This number is only roughly half of the estimated share of the cost of foreign goods in the sum of the wage cost and the cost of foreign goods.

In both equations above, the sum of the coefficients for \(D\log(w_t)\) and \(D\log(p_{ft})\), and for the lagged variables, is far below unity. The hypothesis that the sum is unity is rejected by a Wald test. The homogeneity conditions, i.e. that one can multiply \(p_t, w_t,\) and \(p_{ft}\) by an arbitrary positive number without affecting the relationship, requires that the sum of the coefficients is unity, and is therefore also rejected.

There exists other evidence on the relationship between the variables in Equation (1), which tell a different story. As will be discussed further in Section 2 below, the variables \(\log(p_t), \log(w_t),\) and \(\log(p_{ft})\) are cointegrated which means that there exists a relationship between these variables that can be consistently estimated with OLS. Using data for the same period as above we obtain:

\[
\log(p_t) = -0.0298 + 0.6879 \log(w_t) + 0.3276 \log(p_{ft})
\]

(2)

The usual OLS t-values are very large, but as these statistics are not valid here because the usual OLS estimates of the standard deviations of the coefficients are not valid, we do not report them.

The estimated sum of the coefficients for \(\log(w_t)\) and \(\log(p_{ft})\) is 1.0155, which is very close to unity. \(\beta_1 + \beta_2 = 1\) guarantees that the homogeneity condition, i.e. that

\[
\log(\lambda p_t) = \beta_0 + \beta_1 \log(\lambda w_t) + \beta_2 \log(\lambda p_{ft})
\]

(2')

is valid for arbitrary positive number, \(\lambda\).

The results in Equation (2) are also attractive because the relative size of the coefficients for unit labour cost, and the price of foreign goods, are close to the relative shares of these costs in the sum of wage cost and the cost of foreign goods. This equation therefore indicates that all changes in these costs are eventually passed-through into consumer prices, which is a necessary condition for the stability of the cost shares.

Results from estimations of the long-run relationship between the variables in logs, like the one in Equation (2), have been fairly stable over time, but the results from estimations of the relationship between the same variables in log differences have changed very much. Using data for 1960s – 1990s Guðmundur Guðmundsson successfully estimated equations for the rate of inflation where lagged values of the rate of inflation, the rate of change of the wage rate, and the rate of change of the price of imports, were significant explanatory variables (Guðmundsson, 1990). This changed after 1990 when the rate of inflation in Iceland decreased. The standard deviation of the regressions decreased by two thirds, but the standard deviation of the rate of inflation decreased much more, and therefore \(R^2\) decreased. Guðmundsson (2002), and Guðmundsson (2006) documents these changes, but notes that even if changes in the unit labour cost and the import price explain less of the variation in the rate of inflation, these variables must move in some coherence over time. Guðmundsson (1998) discusses the same model but allows for variable time lags.

Pétursson (2002) estimates a model of price and wage formation using data for the period 1974-1999 and finds coefficients for the wage rate and the import price to be significant in an equation for the rate of inflation. In the inflation equations in different versions of CBI’s macro model, QMM, both
changes in import prices, and in unit labour cost, are significant explanatory variables, except in the most recent version where changes in wage rates are absent (Danielsson et al., 2019).

In this paper we will shed some light on the reasons for the large differences between the results from estimations of equations for logs of the variables, and results from estimations of equations for log differences of the same variables, and some of the reasons why the results of estimations of equations for log differences of the variables have changed over time. We will also evaluate the potential of using an equation in the logs of the variables for forecasting inflation.

Many different economic theories have been proposed to explain the inflationary process: In Wicksell’s theory of inflation in Interest and Prices (1898), and in modern New Keynesian models as in Woodford (2003), the difference between the natural rate of interest and the market rate leads to movements in costs and prices. In Friedman and Schwartz (1963), changes in the money supply affect prices and costs, and in theories of the Phillips-curve (Phillips, 1958, and Friedman, 1968) excess demand in goods and factor markets cause changes in prices and costs. In all these cases the CPI will move with unit labour cost, and the cost of foreign goods. The relationship between these variables will be affected by temporary conditions, expectations, and various conventions that affect different markets and the timing of changes in prices and in wage rates, but it seems intuitive that there has to be some stable relationship between them in the long-run. This relationship is the primary object of this paper.

In this paper we discuss cointegrations between the logs of the CPI, the unit labour cost, and the price of foreign goods, and the usefulness of this information for forecasting in section 2. In section 3 we discuss the relationship between cointegration, average growth of the variables, and the homogeneity condition. In Section 4 we discuss estimation of an equation for log difference of CPI. It is shown how the estimated coefficients are functions of coefficients of correlations and standard deviations of the variables. It is shown that the standard deviations have diverged considerably since the turn of the century, and this has contributed to smaller estimated coefficients for log difference of the unit labour cost, and of the price of foreign goods, the rejection of the homogeneity condition, and to the significance of the constant in these equations. Section 5 discusses how the estimated long-run relationship can be used in forecasting. To illustrate to arguments we use the situations in Iceland in 2011, and in 2015, when the trade-unions and the Confederation of Icelandic Enterprise had reached 2½ and 3½ years settlements. The centralised labour market in Iceland makes possible precise forecasts of future changes in the wage rate based on the centralised settlements. Section 6 discusses estimation of an equation for forecasting log difference of CPI using errors from the estimation of a univariate AR equation in residuals from a cointegrated relationship between the logs of the variables as an error-correction term. Section 7 discusses the forecast performance of the inflation equation in CBI’s macro model, QMM, the equation in Section 6, and some naive forecasts. Section 8 concludes.

2. Cointegration and the long-run relationships

We will restrict our analysis to the period after the large decline in the rate of inflation which took place around 1990. Using data for this period, standard tests for unit roots, such as the Augmented Dickey-Fuller test, show that the variables Log($p_t$), Log($w_t$), and Log($pf_t$) are all integrated. Table 1 shows p-values for Augmented Dickey-Fuller tests for the hypothesis that there is a unit root, using data for the period 1992Q1-2020Q4.
Table 1

P-values for the Augmented Dickey-Fuller test statistic

<table>
<thead>
<tr>
<th>Variable</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(p_t)</td>
<td>0.960</td>
</tr>
<tr>
<td>Log(w_t)</td>
<td>0.973</td>
</tr>
<tr>
<td>Log(pf_t)</td>
<td>0.799</td>
</tr>
<tr>
<td>Dlog(p_t)</td>
<td>0.001</td>
</tr>
<tr>
<td>Dlog(w_t)</td>
<td>0.000</td>
</tr>
<tr>
<td>Dlog(pf_t)</td>
<td>0.000</td>
</tr>
</tbody>
</table>

In no case is the hypothesis that there is a unit root in the log of the level of the data close to being rejected.

The table also shows results from ADF-tests of the hypothesis that log differences of the variables are integrated. It is clearly rejected for all the usual levels of significance.

Using data for the period 2003Q1-2017Q4 Johansens cointegration test does not reject the hypothesis that there are 0, 1, or 2 cointegration equations for the variables Log(p_t), Log(w_t), Log(pf_t), and in the case of one cointegration equation the normalized coefficients are 1.000 for Log(p_t), 0.694 for Log(w_t), and 0.334 for Log(pf_t).

We have also other information to rely on for possible coefficients in a cointegration equation for these variables. As noted above, Statistics Iceland (see www.statice.is) estimates the wage cost in the Icelandic economy at roughly 50% of GDP. The same source estimates imports of goods, excluding alumina, at roughly 25% of GDP before the financial crises. This ratio declined slightly to an average of 23% for the period 2009-2020. If we use 25%, the share of wage cost in the sum of wage cost and cost of foreign goods is 67%, and the share of foreign goods 33%. Ólafsson et al. (2011) report (p. 23) the median cost of imported inputs, and wage cost, in total production costs of firms in their sample. The former is 15% while the latter is 40%, which gives a share of wage cost in the sum of the cost of wages and imported goods of 73%. Pétursson (2002) discusses cointegration between these variables and decides, based on data for the period 1974-1999, to use the coefficients 0.6 for unit labour cost and 0.4 for import prices.

Regressing Log(p_t) on Log(w_t), Log(pf_t), and a constant, using data for the period 1980Q1-2020Q4 gives the following result:

\[
Log(p_t) = -0.0230 + 0.5912 \log(w_t) + 0.3749 \log(pf_t)
\]  
(3)

The same regression using data for the period 2003Q1-2017Q4 was given in Equation (2) above.

The coefficients in Equation (3) (and 2) are consistently estimated if the variables in the equations are cointegrated, i.e. the error terms in the equations are stationary. They are similar, and they are also fairly close to the estimates that can be obtained directly from the national accounts. It is therefore surprising that if we take the difference of the variables in Equation (2’) and estimate the equation in differenced form, using the same data, the estimate of the coefficient for Dlog(w_t) is negative and insignificant.

The error terms in Equations (3) and (2) are highly autocorrelated. Taking the difference of an AR(q) term gives an AR(q+1) term, but, as mentioned in connection with Equations (1) and (1’) above, the
residuals from estimations of the equations in log differences usually pass standard tests for no autocorrelation, and normality of the residuals.

Below we will use 0.65 as a coefficient for the unit labour cost and 0.35 for the price of foreign cost in ISK. Using data for 2003Q1-2017Q4, the Augmented Dickey-Fuller test rejects the hypothesis that \( E65_t = \log(p_t) - 0.65\log(w_t) - 0.35\log(pf_t) \) has a unit root with a p-value of 0.0017. If the data set is extended to 1980Q1-2020Q4 the p-value is 0.0521.

Today, every econometrician knows the dangers of spurious correlations when using regressions to analyse non-stationary data. They also know that these problems can be avoided by suitably differencing the data. But through differencing the data some information, which is embedded in the relationships between the integrated variables, may be lost, as noted in Hamilton (1994) and Harvey (1980). If the variables in the regression are cointegrated, it is possible to use the estimated distribution of the error term to determine the probabilities of possible relations between the variables in the equation. A widely used method in such cases is to include the error term from the estimated regression of the variables in levels in an equation of the variables in differenced form. This is the so-called error-correction methodology due to Davidson et al. (1978). In our case, using data for the period 2003Q1-2017Q4, including lagged values of \( E65_t \) in an equation for the variables in differenced form like Equation (1'), does not give a significant estimate of the coefficient for \( E65_{t-1} \). The p-value is 0.37. If we instead include the residuals from the estimation in Equation (2) above, lagged by one period, the p-value is still higher, or 0.44.

We will discuss the relationship between the equation in logs of the levels, and the equation in log differences further below, but here we note that it is possible to use the information from estimated regressions in log levels in a more direct manner for forecasting. Figure 1 shows values for \( E65_t \) during 36 quarters’ moving windows, the first spanning the periods from 1980Q1-1988Q4 and the last spanning the period 2012Q1-2020Q4.

**Figure 1**

Figure 1 shows that there have been large shifts in upper bounds for \( E65_t \) (Average+2*Standard deviation, and Maximum) and the lower bounds (Average-2*Standard deviation, and Minimum). It also shows that the difference between the upper and the lower bounds have declined over time, mainly due to a shift around the turn of the century. But since then, these bounds have been relatively stable, and could therefore be used in forecasting by assuming that the upper bounds of \( E65_t \) during some recent period can be considered an upper bound for \( E65_t \) in the forecast period, and the lower bounds can be considered a lower bound for \( E65_t \) in the forecast period. Upper and lower bounds for \( \log(p_t) \)
can then be calculated, conditional on the values for \( \text{Log}(w_t) \) and \( \text{Log}(pf_t) \), and used to provide upper and lower bounds for the forecast of the rate of inflation.

Even if the bounds in Figure 1 have narrowed they are still very wide. They might seem relevant for forecasting when inflation is very high as in the 1980s when the average quarterly change was 8.1%, but not now when the quarterly change is below 1% most of the time. But there are two considerations that may make forecasting using the long-run relationship relevant in an era of low inflation. Firstly, if the equation in log levels is used for forecasting the standard deviation of the error of the forecasting the average rate of inflation \( T \) periods ahead decreases faster when \( T \) increases. In that case the standard deviation of the forecast error declines proportionally to \( 1/T \), but if the formula in log differences is used the standard deviation of the error in forecasting average inflation declines proportionally to \( 1/\sqrt{T} \). Secondly, if the value of 0.65Log(\( w_t \)) + 0.35Log(\( pf_t \)) are probable upper bounds for changes in Log(\( p_t \)). In the same way we can calculate a lower bound for the average rate of inflation, conditional on changes in unit labour cost and prices of foreign goods in ISK, if the value of 0.65Log(\( w_t \)) is near the lower bound in the period when the forecast is made. This will be discussed further in Section 5 below in connection with the situations in Iceland in 2011, and in 2015.

3. Cointegration, average growth, and homogeneity

As \( \text{Log}(p_t) \), \( \text{Log}(w_t) \), and \( \text{Log}(pf_t) \) are cointegrated with the cointegrating vector \( [1, -\beta_w, -\beta_pf] \) the following relationship is valid in period \( T \):

\[
\text{Log}(p_T) = \beta_0 + \beta_w \text{Log}(w_T) + \beta_pf \text{Log}(pf_T) + u_T
\]

and also in period 0:

\[
\text{Log}(p_0) = \beta_0 + \beta_w \text{Log}(w_0) + \beta_pf \text{Log}(pf_0) + u_0
\]

Subtracting the second equation from the first, and dividing by \( T \), gives:

\[
\frac{\text{Log}(p_T) - \text{Log}(p_0)}{T} = \beta_w \frac{\text{Log}(w_T) - \text{Log}(w_0)}{T} + \beta_pf \frac{\text{Log}(pf_T) - \text{Log}(pf_0)}{T} + \frac{u_T - u_0}{T}
\]

(4)

As the variables are cointegrated, the residuals, \( u_T \), and \( u_0 \), are bounded, and therefore \( u_T - u_0 \) is bounded. Therefore the last term in (4) can be made arbitrarily small by increasing the value of \( T \). Note also that \( [\text{Log}(x_T) - \text{Log}(x_0)]/T = \overline{\text{Log}}(x_T) \) is the average value of the log difference of the variable \( x \) from period 1 to period \( T \).

Intuitively, it seems reasonable to expect that the average rates of growth of the three variables are equal over long periods of time. It is possible to test these three hypothesis: first if \( \frac{1}{T} [\text{Log}(p_T) - \text{Log}(p_0)] = \frac{1}{T} [\text{Log}(w_T) - \text{Log}(w_0)] \), secondly if \( \frac{1}{T} [\text{Log}(p_T) - \text{Log}(p_0)] = \frac{1}{T} [\text{Log}(pf_T) - \text{Log}(pf_0)] \), and thirdly if \( \frac{1}{T} [\text{Log}(w_T) - \text{Log}(w_0)] = \frac{1}{T} [\text{Log}(pf_T) - \text{Log}(pf_0)] \).

For the Icelandic data discussed above we find that these hypothesis are never rejected given sufficiently long period of time, e.g. the period 1980Q1-2020Q4, or 2003Q1-2017Q4. Given that the last term in Equation (4) is arbitrarily close to zero for sufficiently large value of \( T \), and the means of the log differences of the variables are equal, we can rewrite Equation (4) as:

\[
\mu = \beta_w \mu + \beta_pf \mu
\]
where $\mu$ is the common mean for the log differences of the variables $p_t, w_t,$ and $pf_t$. If $\mu \neq 0$ this equation is valid only if $\beta_w + \beta_{pf} = 1$. We can therefore conclude for these data that the estimated equation for the variables in log levels ensures homogeneity because the average growth of the three series are equal over long periods of time.

Let us finally note that in the case where all variables are growing at the same average rate, all vectors, $[1, -\beta_w, -\beta_{pf}]$, where $\beta_w + \beta_{pf} = 1$, are cointegrating vectors. On the other hand, if the average growth rates of the variables are different, the variables could be cointegrated, but the coefficients might not meet the homogeneity condition.

4. The role of the distribution of the variables in the equation in differenced form, the coefficient estimates, and the homogeneity condition

As noted in the previous section, the estimates of the coefficients for $\log(w_t)$ and $\log(pf_t)$ in Equation (2) above, depend partly on the averages of log differences of the variables. Estimates of the coefficients in the equation in log differences depend, on the other hand, on the correlation between the log differences of the variables and the variation in the log differences. It is possible to write the OLS estimates of the coefficients in the equation

$$D\log(p_t) = \beta_0 + \beta_1 D\log(w_t) + \beta_2 D\log(pf_t) + u_t$$  \hspace{1cm} (5)

which are given in matrix form as $\hat{\beta} = (X'X)^{-1}X'Y$, as follows (see Appendix A for details):

$$\hat{\beta}_1 = \frac{s_{dp}}{s_{dw}} \cdot \frac{r_{dp,dw} - r_{dp,dpf}r_{dw,dpf}}{1 - r_{dw,dpf}^2}$$

$$\hat{\beta}_2 = \frac{s_{dp}}{s_{dpf}} \cdot \frac{r_{dp,dpf} - r_{dp,dw}r_{dw,dpf}}{1 - r_{dw,dpf}^2}$$

and

$$\hat{\beta}_0 = D\log(p) - \hat{\beta}_1 D\log(w) - \hat{\beta}_2 D\log(pf)$$

where $s_{dx}$ is the sample standard deviation of $D\log(x_t)$, and $r_{dx,dy}$ is the sample coefficient of correlation between $D\log(x_t)$ and $D\log(y_t)$.

These formulas show that if $s_{dw}$ and $s_{dpf}$ are large compared to $s_{dp}$, the estimated coefficients, $\hat{\beta}_1$ and $\hat{\beta}_2$, will tend to be low, even in cases where the coefficient of correlation is high.

The figure to the left in the lower row in Figure 2 below shows that since late 1990s 36 quarters estimates of $r_{dw,dpf}$ are close to zero, and almost all of the time slightly negative. In this case replacing $r_{dw,dpf}$ by zero gives reasonable approximations of the coefficients, i.e. $\hat{\beta}_1 \approx \frac{s_{dp}}{s_{dw}} r_{dp,dw}$, and $\hat{\beta}_2 \approx \frac{s_{dp}}{s_{dpf}} r_{dp,dpf}$. As $r_{dp,dw} \leq 1$, and $r_{dp,dpf} \leq 1$, and the standard deviations are positive, we have in this case that

$$\hat{\beta}_1 \approx \frac{s_{dp}}{s_{dw}} r_{dp,dw} \leq \frac{s_{dp}}{s_{dw}}$$

and

$$\hat{\beta}_2 \approx \frac{s_{dp}}{s_{dpf}} r_{dp,dpf} \leq \frac{s_{dp}}{s_{dpf}}$$
The figures in Figure 2 show sample metrics calculated over 36 quarters moving windows. The upper left hand figure shows the standard deviations. In the 1980s and early 1990s they were similar, but in the late 1990s and in the early 2000s they start to diverge. The variation in the exchange rate, and in the price of foreign goods in ISK, increased very much after the fixed exchange rate policy of the 1990s was replaced by the flexible exchange rate policy and inflation targeting in 2001.

**Figure 2**

Moving standard deviations, 36-quarters windows

Moving averages, 36-quarters windows

Moving correlations, 36-quarters windows

Estimated coefficients, 36 quarters windows

If we consider the data for the period 2003Q1-2017Q4 we find that $s_{dp}/s_{dpf} = 0.0112/0.0627 = 0.18$, and that $s_{dp}/s_{dw} = 0.0112/0.0138 = 0.81$. If we take the correlations into account we find that $\hat{\beta}_1$ is actually negative because $r_{dp, dw} < 0$, while $\hat{\beta}_2$ is a fairly low number, mainly because $s_{dpf}$ is so much larger than $s_{dp}$. The sum $\hat{\beta}_1 + \hat{\beta}_2$ is far below unity and the homogeneity condition is rejected.

These results have a simple explanation. As the dependent variable is much smoother than the explanatory variables a least squares algorithm reaches its lowest sum of squared residuals if it is allowed to scale down the effects of the large variability in the right hand side variables to match the lower variability in the variable on the left hand side. The constant becomes significant here because it is smooth and reduces the variation in the outcomes on the right hand side of the equation. Its value is determined so that the average outcomes on the right hand side of Equation (5) are close to the average of the rate of inflation on the left hand side. If we assume that the variables $D\log(p)$, $D\log(w)$, and $D\log(pf)$ have a common population mean, $\mu$, calculating expectations on both sides of Equation (5) gives:

$$\mu = \beta_0 + \beta_1 \mu + \beta_2 \mu \iff \beta_1 + \beta_2 + \beta_0/\mu = 1$$

which has an obvious kinship to the homogeneity condition.

The figure on the right in the lower row in Figure 2 shows moving 36 quarters estimates of the coefficients in Equation (5) above. It shows how the estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ decrease after the end of the high inflation period in the 1980s, and, at the same time, the estimate of the constant, here normalised...
by dividing by the sample average of $D\log(p_t)$, increases. The figure shows that the sum, $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_0/D\log(p_t)$, is close to unity as explained in Equation (6).

Smooth variables, other than the constant, could function in the same way in Equation (6) as does the constant, and their significance may depend more on their role as smoothers of the outcomes on the right hand side, than on their economic function in the inflationary process.

The constant has a mean of unity. If the smooth variable that is added has a mean which is close to the mean of the other variables, and its coefficient is $\hat{\beta}_0$, the homogeneity condition, $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_0 = 1$, will be met, even if the variable is irrelevant for the inflationary process.

The figure in the lower left corner of Figure 2 shows the coefficients of correlation between the pairs of $D\log(p)$, $D\log(w)$, and $D\log(pf)$ over moving 36 quarters periods. It shows that all three coefficients of correlation are high in the 1980s but decline after 1990 when the rate of inflation declines. It is especially the correlation between $D\log(w)$ and the other variables that declines after 1990. The coefficient of correlation between $D\log(w)$ and $D\log(pf)$ becomes negative towards the end of the last century when the data from the pre 1990 high inflation era drop out of the calculations. In the period around the financial crisis of 2008 the coefficients of correlation between $D\log(w)$ on the one hand, and both $D\log(pf)$ and $D\log(p)$ on the other, become negative. This reflects that the financial crises caused moderate increases in the wage rates at the same time as the depreciation of the króna caused large increases in the price of imported goods, and, consequently, in the rate of inflation. The drop in the coefficient of correlations between log changes in the CPI and log changes in unit labour cost when the data for the quarters of the financial crises of 2008 enter the calculations, and the upward jump when these data stop being part of the calculations 9 years later, are very visible in the figure.

5. Forecasting the rate of inflation in 2011 and 2015

Settlements in the Icelandic labour market have frequently prescribed large increases in wage rates, increases well in excess of the sum of the inflation target of 2.5%, and the increase in labour productivity. The trend growth of labour productivity used to be estimated around 2% annually (See CBI, 2021). It has been somewhat lower since the financial crises, which means that during that period the sum of the inflation target and trend growth in labour productivity was below 4.5%.

In May 2011 the representatives for the trade-unions and the Confederation of Icelandic Enterprise agreed on a settlement which prescribed increases in the wage rate, measured by increase in W in CBI’s macro model, QMM, from 2011Q1-2014Q1, of 5.9% per annum. And in 2015, a new settlement prescribed still larger increases in the wage rate. In the period 2015Q1-2018Q4 the average increase in the wage rate amounted to 7.7% per annum. The terms of these contracts were 2½, and 3½ year, which are rather long for Iceland.

When these settlements had been concluded many experts expressed their concerns that the outcomes would lead to spikes in inflation. In the first case the worries of the experts were justified, but in the second case there was no spike.

In this section we will discuss forecasts of average rates of inflation for $t$ periods using the simple formula

$$\frac{1}{t} D\log(p_t/p_0) = \frac{1}{t}[0.65D\log(w_t/w_0) + 0.35D\log(pf_t/pf_0)]$$
The values of the variables at time 0 are known, but their values at time \( t \) are forecasts. When evaluating forecasts with the simple formula we use the forecasts of \( w_t \), and \( p^f_t \), in the baseline forecasts of the CBI.

If we let the true value of the variable \( x \) in period \( t \) be \( x^*_t \), the error in the forecast of the average rate of inflation using the simple formula above can be written as:

\[
\frac{1}{t} [Dlog(p^*_t/p_0) - 0.65Dlog(w_t/w_0) - 0.35Dlog(p^f_t/p^f_0)]
\]

Subtracting and adding \( \frac{0.65}{t} Dlog(w^*_t/w_0) + \frac{0.35}{t} Dlog(p^f_t/p^f_0) \) and rearranging gives the following formula for the error:

\[
\frac{1}{t} [Dlog(p^*_t/p_0) - 0.65Dlog(w_t/w_0) - 0.35Dlog(p^f_t/p^f_0)]
+ \frac{0.65}{t} [Dlog(w^*_t/w_0) - Dlog(w_t/w_0)] + \frac{0.35}{t} [Dlog(p^f_t/p^f_0) - Dlog(p^f_t/p^f_0)]
\]

(7)

The first term in Equation (7) is the forecast error when the true values of the unit labour cost \( (w^*_t) \) and the price of foreign goods in ISK \( (p^f_t) \) are known. This is the error which is due to the simple equation. The second term is the error which is due to the error in the forecast of unit labour cost and the last term is the error which is due to the error in the forecast of price of foreign goods in ISK.

The figure on the left in Figure 3 shows the errors in the forecasts of average quarterly changes. The "formula error" is the first term in Equation (7), "0.65*error in ulct" is the second term, and "0.35*error in pf" is the third term. The last line in the figure is the average error in the baseline forecast for the rate of inflation in the Monetary Bulletin No. 3, 2011.

**Figure 3**

The labour market in Iceland is quite centralised. For that reason CBI's forecasts of changes in the wage rate, based on a given settlement in the labour market, have been fairly accurate, even in the medium term. The forecast published in the Monetary Bulletin No. 4, 2011 predicted the average increase in the wage rate precisely as 5.9% per annum, while the forecast in the Monetary Bulletin No. 3 2011 predicted it as 5.4%. Forecasts of changes in the unit labour cost have been poorer due to difficulties in forecasting changes in labour productivity, and forecasts of changes in the price of foreign goods in ISK have been still poorer due to problems with forecasting the exchange rate.

CBI's forecasts in 2011 of average changes in unit labour cost were not as accurate as the forecasts of average changes in the wage rate. On average, the trend unit labour cost increased by 1.4% per quarter (5.9% per annum) during 2011Q1-2014Q1, while the forecast in the Monetary Bulletin No. 3, 2011 was 0.9% (3.7% per annum). The errors are slightly positive in the beginning, but turn negative in
2012Q2, and remain negative for the rest of the forecast. The errors in the average change in the price of foreign goods are larger and more volatile, but the main reason for the large errors in the forecast of average inflation using the simple formula is the error in the formula itself. These errors are large and positive and show that the firms have passed on the increases in costs into prices at a much slower rate than the formula predicts. The formula errors decline fast and the error in the forecast of average inflation in 2012Q1 is zero. This means that even if the costs are not passed on into prices immediately, they eventually are.

The forecast in the Monetary Bulletin No. 3, 2011, fares better than the forecast using the simple formula, especially in the beginning. The errors in the forecasts of the average rate of inflation from 2011Q1-2012Q2 are equal, but after that the forecast in the Monetary Bulletin is slightly better.

The figure on the right in Figure 3 shows the values of $E65_t = \log(p_t) - 0.65\log(w_t) - 0.35\log(p_{ft})$, using data in the baseline forecasts in the four issues of the Monetary Bulletin published in 2011. The data for the period 2010-2014, published in CBI (2021, are also shown as the true (or the most recent) estimates of $E65_t$. This figure also shows the minimum and the maximum values of $E65_t$ over 36 quarters periods. The minimum (maximum) for 2011Q1 is then the minimum (maximum) value for $E65_t$ during the period 2002Q2-2011Q1, calculated using data from the baseline forecast in Monetary Bulletin No. 1, 2011. The values in 2010 are calculated from actual values for each quarter, but for later years the values are based on the period ending in 2011Q1, and therefore reflect the information available at that time. In the calculations of the minimum, the value in 2008Q4 was considered an outlier and excluded.

These bounds are very wide, and the values of $E65_t$ for the four forecasts in 2011, and for true values of the variables (taken from the forecast in Monetary Bulletin No. 2, 2021) are well inside the boundaries. In this case, considering the boundaries does not seem to add much information that is useful for forecasting.

Figure 4 provides the same information as Figure 3 for the inflation forecast published in Monetary Bulletin No. 3, 2015, after the settlement in the labour market had been concluded. Changes in the wage rate (W in QMM) were very accurately forecasted. For the period 2015Q1-2018Q4 the average increase was 7.7% per annum, but the forecast in Monetary Bulletin No. 3, 2015, predicted 7.2%, and that in Monetary Bulletin No. 4, 2015, predicted 7.9%. The unit labour cost was also reasonably forecasted. The outcome was 1.3% per quarter (5.4% per annum), but the forecast in Monetary Bulletin No. 3, 2015, was 1.4% per quarter (5.8% per annum).

**Figure 4**

The figure on the left shows that there were large positive errors in the forecasting of changes in unit labour cost at the start of the forecast period but they declined quickly. On the other hand, very large errors in the forecasting of average increases in the price of foreign goods in ISK persisted.
throughout the forecast period. In this case, the errors from using the simple formula, are small, and actually negative for some periods, showing that given the correct values for the unit labour cost, and the price of foreign goods, the formula predicts a rate of inflation below the actual rate. The average annual rate of inflation during the period from 2015Q1-2018Q4 was 2.4%, but given the forecast of the unit labour cost, and the price of foreign goods in ISK, in the baseline forecast in Monetary Bulletin No. 3, 2015, the simple formula forecasts an average rate of inflation of 4.4%. In Monetary Bulletin No. 3, 2015, average inflation was forecasted 3.9% annually.

The figure on the right shows the maximum and the minimum for $E65_\tau$ in the most recent 36 quarters as they were in 2014Q1-2014Q4, but from 2015Q1 they show the maximum and the minimum for the 36 quarters period ending in 2015Q1, based on data used in the forecast published Monetary Bulletin No. 1, 2015. The value in 2008Q4 is excluded as an outlier in the calculation of the minimum. As in Figure 3, the bounds are very wide, but here we find three forecasts, which are very close to the lower bound, and one, the forecast published in Monetary Bulletin No. 4, 2015, which actually goes below the lower bound for $E65_\tau$. This indicates that, given the forecasts of the unit labour cost, and the price of foreign goods, in these forecasts, the forecasted rate of inflation was extremely low.

The variable $E65_\tau$ is a measure of the mark-up of prices over the weighted sum of wage cost and cost of foreign goods. It is easy to see that the mark-up that could be calculated using the forecasts of the costs in these forecasts, and actual, ex-post, consumer prices during the forecast period, would be well below the lower bound, showing that, given the forecast of the unit labour cost, and the price of foreign goods, the actual development of the price level was extremely unlikely.

Figure 4 also shows $E65_\tau$ calculated from the actual values of the variables (labeled MB21-2). The actual mark-up, measured by $E65_\tau$, is much higher than the forecasted mark-ups due to the large appreciation of the Icelandic króna, not anticipated by any of the forecasts.

These results do not preclude that other factors like changes in expectations, or improvements in monetary policy, can affect the inflationary process as discussed in Pétursson (2020). It only means that in so far as these factors may influence the process, their effects must be reflected somehow in the interaction between the unit labour cost, the price of imported goods, the CPI, and the restrictions that these variables must satisfy. As the wage rate was accurately forecasted in the forecasts published in the latter half of 2015 – and wage rates are unlikely to change much in Iceland during the term of a contract, the factors that would moderate the rate of inflation, compared to the rate predicted by the simple formula, would have to work through the large unexpected change in the price of foreign goods in this period, which means that they have to work through the large appreciation of the exchange rate.

Let us finally take a quick look at recent developments in the mark-up. Figure 5 shows the mark-up measured by $E65_\tau$ in the period 2018Q1-2020Q4, based on data used in the forecast published in Monetary Bulletin No. 2, 2021.
The figure shows that the mark-up, measured by $E65_t$, declined in 2018, and approached the lower bounds in early 2019. It increased again in late 2019 when productivity increased, and unit labour cost decreased at the same time as the the kröna appreciated. It declined in 2020 when decline in productivity due to the Covid-19 pandemic increased trend unit labour cost. The year 2020 should probably be treated as an outlier for the relationships discussed in this section, but the fact that $E65_t$ is, and has been, close to, and sometimes below, the lower bound, is relevant information for analysis of future developments of the rate of inflation when the economic conditions become normal again.

6. Estimating an equation for inflation with error-correction

It was noted in Section 2 above that if the error term from the estimation of the equation for the variables in logs, lagged one period, or the variable $E65_t$, lagged one period, is included in an equation for the rate of inflation, the coefficient is not significant when the data period is 2003Q1-2017Q4. We have estimated equations where the coefficient for $E65_t$ has been significant, and negative, but we have found that the estimated equation can be improved (measured by $R^2$ (adj.) or Akaike info criterion) by using the error from a univariate equation for the variable $E65_t$. Using data for the period 2003Q1-2017Q4 that were used in the forecast published in Monetary Bulletin No. 2, 2021, we obtain the following equation (t-values in parenthesis):

$$E65_t = -0.02199 + 0.56108 E65_{t-1} - 0.2262 E65_{t-4}$$

R$^2$ for this equation is 0.36, standard error of the regression is 0.015, and appropriate tests accept the hypothesis that the residuals are normal, homoskedastic and without autocorrelation. The difference equation in Equation (8) is stable, and if there are no shocks it will return to the equilibrium value of $E65_t = -0.0329$.

It is now possible to use the estimated error term of Equation (8), lagged one period, as an explanatory variable in an equation which also contains variables that have been used in the inflation equations in the various versions of CBI’s QMM model. Equation (9) below is obtained using data for the period 2003Q1-2017Q4. The data were used in the forecast published in Monetary Bulletin No. 2, 2021, except for the output gap where we use the data that were used in the estimation of the equation for inflation in QMM 4.0.
\[
D\log(p_t) = 0.2374D\log(p_{t-1}) + 0.1236(d\log(w_t) - d\log(w_{t-1})) + 0.0901D\log(pf_t)
\]
\[
-0.2364[E65_{t-1} - 0.5611E65_{t-2} + 0.2262E65_{t-5}] + 0.3991(gapav_{t-1} - gapav_{t-2})
\]
\[
+0.1005gapav_{t-4} + 0.0151D082 - 0.0153D032 - 0.0042Q1 + 0.0035Q2 - 0.0054Q3
\]
\[
+0.0039 \cdot S1202
\]

In Equation (9) Q1, Q2, and Q3 are centered seasonal dummies, D082 is a dummy taking the value 1 in 2008Q2, but zero otherwise, D032 is a dummy taking the value 1 in 2003Q2 but zero otherwise, and S1202 is a dummy taking the value 1 in quarters up to an including 2012Q2, but zero after that. S1202 is closely related to \(Log(1 + SIT \cdot 0.025)/4\), an index measuring confidence in CBI’s monetary policy defined in Pétursson (2020). If this index is included in Equation (9) instead of S1202 the estimated coefficient is highly significant, but overall, Equation (9) is slightly better with S1202. This equation has an \(R^2\) of 0.900, and the standard error of the regression is 0.0039. It passes usual tests for normality, homoskedasticity and no autocorrelation of the residuals.

In Equation (9) the coefficient for the lagged one quarter change in the log difference of unit labour cost is significant. The unit labour cost affects the rate of inflation also through the error-correction term, which ensures that in the long-run all increases in the unit labour cost, and in the price of foreign goods, are passed through into prices, but in the short-run of two quarters, estimated pass through of changes in the unit labour cost is zero. The estimated pass through of changes in import prices is significant and positive, but only a small part of the long-run pass through determined by the error correction term.

7. Comparing out-of-sample inflation forecasts using different formulas

Table 2 below presents Root Mean Square Errors (RMSEs) for out of sample one step ahead forecasts using different forecast equations. In all cases we use the data used in the forecast presented in Monetary Bulletin No. 2, 2021. We consider the following methods for forecasting inflation: 1) The inflation equation in the present version (4.0) of CBI’s macro model, QMM; 2) Equation (9) in Section 6 above; 3) The forecast is the most recent value for log difference of CPI; 4) The forecast is the average of the four most recent values of Dlog(CPI); and 5) The forecaster has full confidence in CBI’s ability to keep the inflation at the 2.5% target, and assumes all deviations from the target to be random. This forecaster rationally forecasts in all cases that the quarterly change in the log of CPI will be \(Log(1.025)/4 = 0.006173\). The three last forecast equations are so called naive forecasts which should generally be poorer than the more sophisticated ones.

<table>
<thead>
<tr>
<th>Year Q1-Q2</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018Q1-2020Q4</td>
<td>0.00550</td>
<td>0.00530</td>
<td>0.00597</td>
<td>0.00429</td>
<td>0.00403</td>
</tr>
<tr>
<td>2018Q1-2020Q4, 2020Q2 excl.</td>
<td>0.00568</td>
<td>0.00497</td>
<td>0.00461</td>
<td>0.00349</td>
<td>0.00338</td>
</tr>
<tr>
<td>2018Q1-2019Q4</td>
<td>0.00546</td>
<td>0.00487</td>
<td>0.00480</td>
<td>0.00320</td>
<td>0.00285</td>
</tr>
</tbody>
</table>

Table 2 shows the RMSEs for the period 2018Q1-2020Q4. It also shows the RMSEs for this period when 2020Q2 is excluded, as this one quarter affects significantly the performance of the naive
forecasts. It also seemed reasonable to show the RMSEs for the period before the Covid-19 pandemic started to affect the economy.

Table 2 shows that forecasting all the time that $Dlog(p_t)$ is 0.006173 gives lowest RMSE in all cases. Forecasting using the average of the four most recent values of $Dlog(p_t)$ is second best. Forecasting using the most recent value of $Dlog(p_t)$ has much higher RMSEs than forecasting with the average of the four most recent values. The RMSEs for Equation (9) in Section 6 are much higher than the RMSEs of the best naive forecasts, but marginally lower than the RMSEs for the inflation equation in CBI’s QMM for all periods shown in Table 2.

8. Conclusions

We showed in Section 4 how the estimated coefficients in the regression of $Dlog(p_t)$ on $Dlog(w_t)$, $Dlog(pf_t)$, and a constant, depend on the coefficients of correlations between these variables and their standard deviations. Larger variation in $Dlog(w_t)$, and especially in $Dlog(pf_t)$, than in $Dlog(p_t)$, contribute to lowering the estimates of the coefficients for log difference of unit labour cost, and for log difference of the price of foreign goods, and therefore also to the rejection of the homogeneity condition that the sum of these coefficients is unity. As the large variations in $Dlog(pf_t)$ dominate movements in $Dlog(p_t)$ they also lead to low, and sometimes negative, estimates of the coefficient of correlations between $Dlog(w_t)$ and $Dlog(p_t)$. Economic crises in Iceland are frequently characterized by low increases in wage rates, combined with large depreciations of the króna that fuel increases in the price of foreign goods, and in the rate of inflation.

Large sample standard deviations for $Dlog(w_t)$, and $Dlog(pf_t)$, compared to the sample standard deviation for $Dlog(p_t)$, lead to low values of the estimates of the coefficients for $Dlog(w_t)$, and $Dlog(pf_t)$. As the expected values of $Dlog(w_t)$, $Dlog(pf_t)$, and $Dlog(p_t)$ are equal, these low estimates of the coefficients for the right hand side variables create a role for the constant to make the average on the right-hand side equal to the average on the left-hand side, at the same time as it smooths the outcomes on the right hand side. Other variables with small variations could function in the same way as the constant. If the average of this variable is the same as the averages of $Dlog(p_t)$, and of the other variables on the right hand side, the sum of the three coefficients for the variables on the right hand side will be unity, even if the added variable has no significant economic role in the inflationary process, like the constant.

Even if the variable for the log difference of unit labour cost, $Dlog(w_t)$, does not have a significant coefficient in a regression using log differences of the variables, it does not seem wise to ignore possible effects of changes in the price of such a large part of total cost, as the wage cost is, on the rate of inflation. In that case the estimated long-run relationship may be of some help in avoiding errors that may follow from underestimating the influences of changes in unit labour cost on the rate of inflation.

Frequently, including estimated errors from the long-run relationship between the logs of the variables, lagged by one quarter, does not give a significant coefficient for the error-correction term. In our experience, estimating an AR-equation for the errors from the long-run relationship for the logs of the variables, and including the error from that equation, lagged one period, in an equation with other stationary explanatory variables, does give a significant coefficient, with a correct sign, for this more complicated error-correction term. The resulting equation performed better than the inflation equation in CBI’s QMM in out-of-sample one step ahead forecasts, but not as well as the forecast based on full confidence in CBI’s ability to maintain inflation at the 2.5% target. That indicates that the
smoothing of the effects of changes in the right hand side variables, embodied in Equation (9) in Section 6, are insufficient to match the low variability of the rate of inflation – at least during the period 2018Q1-2020Q4. Possibly, the equation, as it stands, with a variable for the downward shift in average rate of inflation in 2012, does not provide sufficiently for the smoothing effects of the public’s confidence in the CBI’s monetary policy on the inflationary process.

Appendix A

We use \( dp_t \) for \( Dlog(p_t) \), \( dw_t \) for \( Dlog(w_t) \) and \( dpf_t \) for \( Dlog(pf_t) \) and rewrite Equation (5) as:

\[
dp_t = \beta_0 + \beta_1 dw_t + \beta_2 dpf_t + u_t
\]  
(A.1)

Taking average on both sides gives:

\[
\overline{dp} = \beta_0 + \beta_1 \overline{dw} + \beta_2 \overline{dpf} + \bar{u}
\]  
(A.2)

Subtracting Equation (A.2) from Equation (A.1) gives:

\[
dp_t - \overline{dp} = \beta_1 (dw_t - \overline{dw}) + \beta_2 (dpf_t - \overline{dpf}) + u_t - \bar{u}
\]  
(A.3)

The data matrices are:

\[
Y = \begin{bmatrix} dp_t - \overline{dp} \\ \vdots \\ dp_T - \overline{dp} \end{bmatrix}, \text{ and } X = \begin{bmatrix} dw_1 - \overline{dw} & dpf_1 - \overline{dpf} \\ \vdots & \vdots \\ dw_T - \overline{dw} & dpf_T - \overline{dpf} \end{bmatrix}
\]

It follows that

\[
X'X = \begin{bmatrix} \sum (dw_i - \overline{dw})^2 & \sum (dw_i - \overline{dw})(dpf_i - \overline{dpf}) \\ \sum (dw_i - \overline{dw})(dpf_i - \overline{dpf}) & \sum (dpf_i - \overline{dpf})^2 \end{bmatrix}
\]

and

\[
(X'X)^{-1} = \frac{1}{|X'X|} \begin{bmatrix} \sum (dpf_i - \overline{dpf})^2 & -\sum (dw_i - \overline{dw})(dpf_i - \overline{dpf}) \\ -\sum (dw_i - \overline{dw})(dpf_i - \overline{dpf}) & \sum (dw_i - \overline{dw})^2 \end{bmatrix}
\]

Using that the sample variance for \( dx \) is \( s_{dx}^2 = \sum (dx_i - \overline{dx})^2/(T - 1) \), the sample covariance for \( dx \) and \( dy \) is \( Cov(dx, dy) = \sum (dx_i - \overline{dx})(dy_i - \overline{dy})/(T - 1) \), and the sample coefficient of correlation for \( dx \) and \( dy \) is \( r_{dx,dy} = Cov(dx, dy)/(s_{dx}s_{dy}) \), we get that

\[
(X'X)^{-1} = \frac{1}{|X'X|} \begin{bmatrix} (T - 1)s_{dpf}^2 & -(T - 1)s_{dw} s_{dpf} r_{dw,dpf} \\ -(T - 1)s_{dw} s_{dpf} r_{dw,dpf} & (T - 1)s_{dw}^2 \end{bmatrix}
\]

\[
= \frac{(T - 1)^2 s_{dpf}^2 s_{dw}^2}{s_{dpf}^2 s_{dw}^2 (1 - r_{dw,dpf}^2)} \begin{bmatrix} s_{dpf}^2 & -s_{dpf} s_{dw} r_{dw,dpf} \\ -s_{dpf} s_{dw} r_{dw,dpf} & s_{dw}^2 \end{bmatrix}
\]

where \(|X'X| = \sum (dw_i - \overline{dw})^2 \sum (dpf_i - \overline{dpf})^2 - (\sum (dw_i - \overline{dw})(dpf_i - \overline{dpf}))^2 = (T - 1)s_{dw}^2 (T - 1)s_{dpf}^2 - (T - 1)^2 s_{dpf}^2 s_{dw}^2 r_{dw,dpf}^2 = (T - 1)^2 s_{dpf}^2 s_{dw}^2 (1 - r_{dw,dpf}^2)\)

We have also that
\[
X'Y = \left[ \sum (dp_i - \bar{dp})(dw_i - \bar{dw}) \right] = (T - 1) \left[ \frac{sdp}{sdpsdp} \right] \]

The estimated coefficients are:

\[
\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X'X)^{-1}X'Y
\]

\[
= \frac{(T - 1)(T - 1)}{(T - 1)^2 s_{dp}^2 s_{dw}^2 (1 - r_{dw,dpf}^2)} \left[ \begin{array}{c} s_{dpf} s_{dp} s_{dw} r_{dp, dw} - s_{dw} s_{dpf} r_{dw, dpf} s_{dp} s_{dpf} r_{dp, dpf} \\ -s_{dw} s_{dpf} r_{dw, dpf} s_{dp} s_{dpf} r_{dp, dpf} + s_{dp}^2 s_{dpf} r_{dp, dpf} \end{array} \right]
\]

and therefore:

\[
\hat{\beta}_1 = \frac{s_{dp} s_{dpf} r_{dp, dw} - s_{dpf} r_{dw, dpf} r_{dp, dpf}}{s_{dpf} s_{dw} (1 - r_{dw, dpf}^2)} = \frac{s_{dp} r_{dp, dw} - r_{dw, dpf} r_{dp, dpf}}{s_{dw} (1 - r_{dw, dpf}^2)}
\]

\[
\hat{\beta}_2 = \frac{s_{dp} s_{dpf} r_{dw, dpw} + s_{dw} r_{dp, dpf}}{s_{dpf} s_{dw} (1 - r_{dw, dpf}^2)} = \frac{s_{dp} r_{dp, dpf} - r_{dw, dpf} r_{dp, dw}}{s_{dpf} (1 - r_{dw, dpf}^2)}
\]

Finally, Equation (A.2) gives that:

\[
\hat{\beta}_0 = \bar{dp} - \hat{\beta}_1 \bar{dw} - \hat{\beta}_2 \bar{dpf}
\]

References


WORKING PAPER
2021|87