CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor thesis

2023 **David Červený**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

PPI and CPI: What is the relationship?

Bachelor thesis

Author: David Červený

Study programme: Economics

Supervisor: Mgr. Petr Polák, M.Sc., Ph.D.

Academic Year: 2022/2023

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.
2. The author hereby declares that all the sources and literature used have been properly cited.
3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague, 31st December 2022 David Červený

Bibliographic note

ČERVENÝ, David. *PPI and CPI: What is the relationship?*. Prague, 2023. 65 p. Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor Mgr. Petr Polák, M.Sc., Ph.D.

Length of the thesis: 97611

Abstract

This bachelor thesis examines the relationship between the PPI and the CPI in the Czech Republic and the euro area. The primary method used in this thesis is the Granger causality test. Granger causality between the price indices is tested for in a bivariate model and also conditional on other variables describing the development of real GDP, a given monetary aggregate and wages. The most apparent conclusion that can be drawn from the empirical results indicates that the PPI Granger-causes the CPI in the Czech Republic and that there is no Granger causality going from the CPI to the PPI in the euro area. These results are consistent with conventional economic theory, which suggests a pass-through effect in the production chain going from producer prices to consumer prices.

Keywords

Inflation, Producer price index, Consumer price index, Harmonised index of consumer prices, PPI, CPI, HICP, Euro area, Granger causality

Abstrakt

Tato bakalářská práce se zabývá vztahem mezi indexem cen výrobců a indexem spotřebitelských cen v České republice a v eurozóně. Základní metodou použitou v této práci je Grangerův test kauzality. Grangerova kauzalita mezi cenovými indexy je testována v dvourozměrném modelu a také podmíněně na dalších proměnných popisujících vývoj reálného HDP, daného peněžního agregátu a mezd. Nejzřetelnější závěr, který lze z empirických výsledků vyvodit, naznačuje, že PPI je v České republice Grangerovou příčinou CPI a že v eurozóně neexistuje Grangerova kauzalita směřující od CPI k PPI. Tyto výsledky jsou v souladu s konvenční ekonomickou teorií, která předpokládá průchozí efekt ve výrobním řetězci směřující od cen výrobců ke spotřebitelským cenám.

Klíčová slova

Inflace, index cen výrobců, index spotřebitelských cen, harmonizovaný index spotřebitelských cen, PPI, CPI, HICP, eurozóna, Grangerova kauzalita

Název práce

PPI a CPI: Jaký je vztah?

Acknowledgements

I would like to express my gratitude to Mgr. Petr Polák, M.Sc., Ph.D. for consultation of the topic and for his helpful and constructive advice. I am also grateful to my family for their support.

Contents

List of tables

List of figures

Acronyms

ADF	Augmented Dickey-Fuller
AIC	Akaike information criterion
AR	Autoregression
ARIMA	Autoregressive integrated moving average
BLS	(United States) Bureau of Labor Statistics
BVAR	Bayesian vector autoregression
CCI	Construction cost index
CNB	Czech National Bank
COICOP	Classification of Individual Consumption According to Purpose
CPI	Consumer price index
CZSO	Czech Statistical Office
DF	Dickey-Fuller
ECB	European Central Bank
ECOICOP	European Classification of Individual Consumption According to Purpose
EMU	Economic and monetary union
G7	Group of Seven
GDP	Gross domestic product
HQIC	Hannan-Quinn information criterion
I(0)	Integrated of order 0
I(1)	Integrated of order 1
IQI	Implicit quality index
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LCI	Labour cost index

NACE	Statistical Classification of Economic Activities in the European Community (Nomenclature générale des activités économiques dans les Communautés européennes)
OLS	Ordinary least squares
PPI	Producer price index
PPP	Purchasing power parity
SIC	Schwarz information criterion
SPPI	Services producer price index
SRI	Standard reference index
TVEC	Threshold vector error correction
USD	United States dollar
VEC	Vector error correction
VAR	Vector autoregression
WPI	Wholesale price index

1 Introduction

Inflation is certainly among the most important macroeconomic variables. In a standard way, consumer price inflation is measured by the year-over-year growth of the consumer price index (CPI), while producer price inflation is measured by the year-over-year growth of the producer price index (PPI).

This thesis examines the development of these price indices. The approach of this thesis entails estimation of several OLS regression models as a part of the Granger causality test. Granger causality and conditional Granger causality, both very similar concepts, together create the framework in which the direction of the relationship between the two price indices may be determined. This thesis is concerned with examination of the relationship for the Czech Republic and for the euro area. The Granger causality test regressions are estimated many times because they are sensitive to lag length and also each regression model can only be used to test for one direction of the relationship. The null hypothesis in each one of these tests is that of no Granger causality.

The literature appears far from settled in its entirety, but there is a slight tendency towards Granger causality from the PPI to the CPI. This direction of the relationship would be in concurrence with generally accepted belief, as the CPI may on the whole cover economic activity further down the production chain than the PPI. Any price changes for example in inputs would therefore at first manifest themselves in the PPI and only then in the CPI. Similar theoretical reasoning is provided among others by Clark (1995), who presents quite a characteristic conclusion: there is indeed Granger causality from the PPI to the CPI, but the strength of this relationship is unstable in time. Methodically, Holub (2000) and particularly Fan et al. (2009) are very close to the way the analysis in this thesis is conducted. This thesis also follows Clark (1995) regarding the selection of variables for the multivariate models.

The research question of this thesis is motivated by the possibility of additional clarification of the relationship between the PPI and the CPI using newer data. The relationship also appears to be region-specific and there has not been much Czech research on this topic or even much research using specifically common data for the euro area. If there was a strong and clear conclusion regarding the relationship between the PPI and the CPI, it might for instance be used to improve inflation forecasting. This would certainly be quite a welcome progression for any financial or other institution that relies on inflation forecasting.

The thesis is organized as follows. Chapter 2 serves as a summary of the topic of price indices from a theoretical standpoint and also focuses on some issues that price indices may exhibit. Chapter 3 provides a review of literature on the research question itself. Chapter 4 presents the data and the methods of analysis used in this thesis, while results are presented in Chapter 5 together with their discussion and an important note on the limitations of using the Granger causality framework. Chapter 6 serves as a conclusion.

2 Theoretical background on price indices

Inflation is defined as an increase in the general price level in an economy. Inflation rate is the annual rate of growth of the price level. Price indices, which are the main topic of this thesis, are the means of identifying changes in the price level. As there are many price indices and also several distinct price index formulas, the inflation rate may vary depending on the index expressing it.

The most notable price indices are the CPI and its EU-wide version named HICP (harmonised index of consumer prices), the PPI and the GDP deflator. These indices are further individually described in more detail in this chapter. Less frequently used price indices are often more narrow in their scope or focus only on a particular market in the economy. In the EU, they include for example import and export price indices, the labour cost index (LCI), the construction cost index (CCI) or the services producer price index (SPPI) (Eurostat, 2008). The process of calculating the PPP is also akin to the calculation of price indices.

There are two main formulas for price index construction and calculation – Laspeyres and Paasche. These two formulas are quite similar in their appearance, as they both include ratios of sums of prices, but are distinct in regards to the quantities of goods and services included in the index, which constitute the weights assigned to these prices; the Laspeyres formula uses fixed quantities from the base period and the Paasche formula uses quantities from the current period instead. Other less frequently used price indices calculated by their own methods and formulas include the Törnqvist index, the Marshall-Edgeworth index or the Walsh index. The latter two are relatively similar to the Laspeyres and Paasche formulas, with the Marshall-Edgeworth formula using the arithmetic mean of the base period and the current period quantities as weights and the Walsh formula using the geometric mean of these quantities instead (Anghelache, Manole, Diaconu, Gheorghe, & Ciocan, 2012). Of the most common price indices, the CPI, the HICP and the PPI are Laspeyres indices, while the GDP deflator is a Paasche index.

A Laspeyres index tends to overstate inflation (Mankiw, 2010). This phenomenon, along with an overview of relevant papers, is discussed in greater detail in the CPI section of this chapter. However, neither Laspeyres index nor Paasche index is a clearly superior measure of change in the cost of living because, contrastingly, a Paasche index tends to understate inflation (Mankiw, 2010). A middle approach to tackling this issue is taken by an index formula combining the two into one – Fisher index, also known as Fisher ideal price index, which is the geometric mean of a Laspeyres index such as the CPI and a Paasche index such as the GDP deflator. Diewert (1992) summarizes twenty desirable mathematical properties of price indices, which have been proposed by various researchers over the course of the 20th century, and concludes that the Fisher price index is the only index formula that satisfies all of these twenty properties.

2.1 CPI

The CPI is a price index measuring the average change over time in the prices paid by consumers for a consumption basket of goods and services. It is the most widely used price index (U.S. Bureau of Labor Statistics, 2022). It is an official measure of consumer price inflation in a given country. The CPI is a monthly statistic as are most major price indices.

A price index with a fixed consumption basket of goods and services originating in the base period is called a Laspeyres index. The CPI is an example of a Laspeyres index. Using a similar notation style to Garín et al. (2018), the value of the CPI in period t is equal to the ratio of the cost of the base period (b) consumption basket in period t to the cost of the same basket in the base period itself. All values for quantity (q_i, $i \in \{1, \dots, n\}$) of products are taken from the base period and do not change because the basket is fixed. The cost of the consumption basket in each time period (b, t) is equal to the sum of products of price (p_i, $i \in \{1, \dots, n\}$) and quantity of each good or service included in the CPI calculation. Then, the formula

$$CPI_t = \frac{p_{1,t}q_{1,\mathrm{b}} + p_{2,t}q_{2,\mathrm{b}} + \cdots + p_{n,t}q_{n,b}}{p_{1,b}q_{1,\mathrm{b}} + p_{2,b}q_{2,\mathrm{b}} + \cdots + p_{n,b}q_{n,b}} = \frac{\sum_{i=1}^{n} p_{i,t}q_{i,b}}{\sum_{i=1}^{n} p_{i,b}q_{i,b}}$$

is the expression of this CPI construction in mathematical terms. The consumption basket is supposed to be as representative of consumer behavior in the country as possible. Because of that, the quantities are particularly important as they effectively weight the products and product categories contained in the consumption basket by their share of overall consumer expenditure. The CPI can also be understood as a ratio of weighted averages of prices.

There may be many levels of aggregation of product categories in the process of constructing the overall index. A detailed example is provided in a section dealing with the HICP, an EU-wide version of the CPI. Diligent weighting of different product groups in the index is important for accuracy of the index. However, errors in the price level have a larger effect on the index than errors of the same scale regarding weighting (Eurostat, 2018).

National statistical offices or similar institutions process data in order to determine the composition of a representative consumption basket as accurately as possible and to determine new price levels of this consumption basket in subsequent periods. An important modification of the index named core CPI measures the price change of a narrowed consumption basket, from which prices of food and energy are excluded. Prices in these consumption segments tend to be quite volatile in the short term and hence core CPI may be perceived as a better inflation indicator than the CPI itself (Mankiw, 2010).

The CPI has been in use in the United States since 1917 and its monthly calculation and publication is the responsibility of the Bureau of Labor Statistics. It is composed of prices in the following main groups of goods and services: Food and beverages, Housing, Apparel, Transportation, Medical care, Recreation, Education and communication, Other goods and services (Samuelson & Nordhaus, 2009). Laws, private contracts or government transfer payments such as Social Security are indexed to inflation through cost-of-living allowances, which use the CPI to adjust for changes in the price level (Mankiw, 2010).

Interestingly, while the BLS defines the CPI as "a measure of the average change overtime in the prices paid by urban consumers for a market basket of consumer goods and services" (U.S. Bureau of Labor Statistics, 2022), other countries and entities such as the EU or the Czech Republic do not define the CPI as representative of urban consumers and instead pursue geographical weighting of urban and rural outlets to ensure representativeness of the consumption basket.

There are also some issues associated with the CPI. Of these issues, substitution bias is one of the more prominent ones. Resulting from being a Laspeyres index, the CPI tends to overstate increases in the cost of living because the fixed consumption basket from the base period does not take into account the possibility of consumers substituting goods and services in response to changes in their relative prices (Lebow & Rudd, 2003). Besides, the traditional framework of the CPI has been a cost-of-goods index rather than a cost-of-living index (Schultze, 2003). Although a cost-of-living approach is also taken into consideration while constructing the CPI (U.S. Bureau of Labor Statistics, 2022), a true cost-of-living index would measure the consumer expenditure needed to maintain a given level of utility (Lebow & Rudd, 2003). The Laspeyres index formula does not take utility levels into consideration, as the measurement of those would arguably be complicated.

If the CPI considerably overstates the cost of living and displays a high upward bias, the consequences may be important for fiscal policy because various programs of government spending are indexed to it. Furthermore, an upward bias in the CPI leads to a downward bias in the measurement of GDP growth (Groshen, Moyer, Aizcorbe, Bradley, & Friedman, 2017). Groshen et al. (2017) estimate that the bias within real GDP growth in the United States introduced by mismeasurement of CPI segments, notably medical care, prescription drugs and information and communications technology, was −0.2 percentage points in 2000 and −0.26 percentage points in 2015. Greenstein and McDevitt (2011) focus on the category of internet access services and find that a sub-index for internet access services made by the BLS has exhibited slower price decreases than their own such index possibly due to inadequate quality adjustment of the prices over time. This result supports the notion of upward bias in the information and communications technology segment of the CPI.

Shapiro and Wilcox (1996) identify several effects that are sources of bias in CPI measurement in the United States. Apart from across-strata and within-strata effects, both of which are forms of the already mentioned substitution bias, they recognize new-items effect, new-outlets effect and quality change effect. Each of these effects creates bias within the CPI with estimates ranging from 0.1 to 0.25 percentage points per year (Shapiro & Wilcox, 1996). The individual biases and the overall bias are also each assigned a subjective probability distribution by the authors. The overall bias is perceived as a random variable obtained by aggregating the biases resulting from all of the individual effects. The median of the subjective probability distribution of the overall bias is just under 1 percentage point per year with an 80 % confidence interval ranging from 0.6 to 1.5 percentage points per year (Shapiro & Wilcox, 1996).

Also in 1996, the *Advisory Commission To Study The Consumer Price Index* chaired by economist Michael Boskin presented a report on bias in the computation of the CPI to the Senate Finance Committee. The *Advisory Commission* concluded that the CPI was biased upward by 0.8 to 1.6 percentage points per year with a point estimate of 1.1 percentage points per year (Mankiw, 2010). These findings are similar to those of Shapiro and Wilcox.

Lebow and Rudd (2003) revisit the issue of CPI mismeasurement in the United States. The authors deem inadequate accounting for quality improvements and the introduction of new items to be the largest sources of bias within the CPI. Their study distinguishes among the same sources of bias as the previous research done by Shapiro and Wilcox with an addition of weighting bias, which was not previously discussed. Each of the effects creates bias within the CPI with estimates ranging from 0.05 to 0.37 percentage points per year and the upper bound value of 0.37 percentage points per year is an estimate for the bias caused simultaneously by the new-items effect and the quality change effect, which together form the most uncertain components of bias in the CPI (Lebow & Rudd, 2003). This uncertainty is reflected by the broad range of the 90 % confidence interval assigned by the authors to the estimate of the bias resulting from these two effects, which ranges from −0.1 to 0.8 percentage points per year. The overall bias estimate is lower than in previous studies and in the authors' opinion it is the case due to improvements in the BLS methodology. The overall bias estimate is approximately 0.6 percentage points per year with a 90 % confidence interval ranging from 0.1 to 1.2 percentage points per year (Lebow & Rudd, 2003).

The issues of quality change in the basket of goods and services and subsequent quality adjustment of the index are complicated ones but methodology has been focused on them. More detail about this topic is provided in the section about the HICP since the quality change effect may potentially be a source of bias in CPI measurement in the EU as well.

However, there is also some conflicting evidence regarding the claim of upward quality change bias in the CPI. Gordon and vanGoethem (2003) explore the possibility of downward bias within the rental housing component of the CPI in the United States throughout its entire history of measurement. They claim that this effect may have been particularly pronounced during periods when data on rents have been used to impute price changes for owner-occupied housing and that it places uncertainty on the common conclusion that the overall CPI is biased upward. Moreover, Kryvtsov (2016) uses a price survey dataset compiled by Statistics Canada, which consists of prices for goods and services in Canadian retail outlets in the period 1998–2006, to show that the quality change effect has not been a notable source of any bias in the measurement of the CPI in Canada.

2.2 HICP

As this thesis is focused primarily on the EU and the euro area, this part covers the practice of CPI calculation in the EU in more detail. EU member states use a specific form of the CPI – the HICP. The HICP offers comparability of consumer price inflation among countries of the euro area and the EU.

The HICP for individual EU member states is calculated by their national statistical offices and the aggregate HICP for groups of countries such as the euro area is produced by the Eurostat. The purpose of the HICP is to be primarily a macroeconomic policy indicator. It is used to define price stability, which is targeted by the ECB and other central banks, and it is also used to determine price convergence between countries preparing for accession to the euro area and the euro area itself.

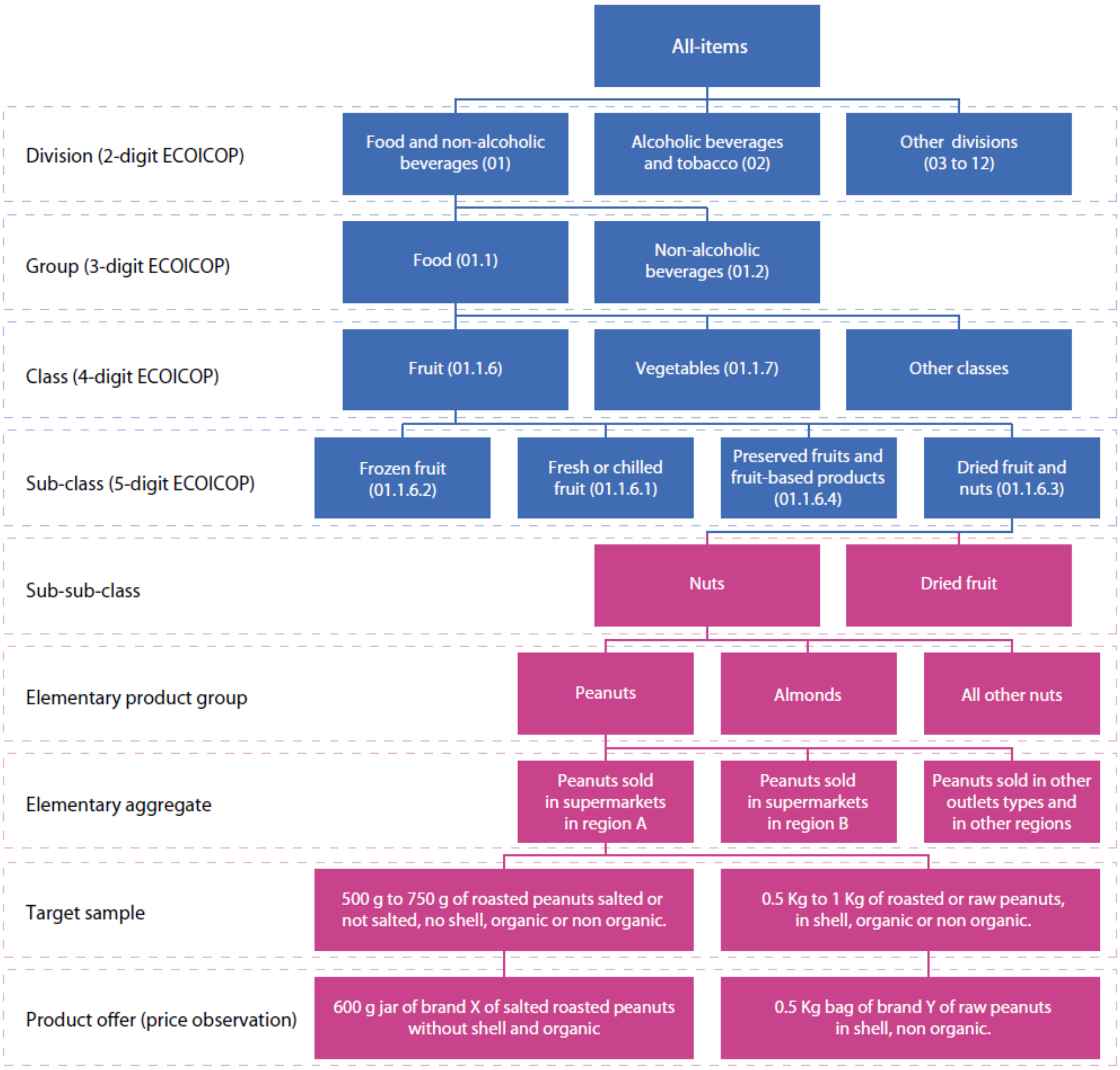
The HICP has been in use since 1997. Like other price indices, it is a monthly statistic and data for its calculation is collected every month. Eurostat (2018) describes the beginnings of the HICP in more detail as follows.

> "Countries were either unwilling or unable to change a key statistic like the national CPI to serve the then limited purpose of international comparison.

> This situation changed in 1992, when the Maastricht Treaty laid down criteria for joining the euro area (i.e. stage III of economic and monetary union (EMU)). One such criterion was sustainable convergence in price stability, to be assessed in comparison with the best performing Member States.
>
> So, with the Maastricht Treaty, it became imperative to compare consumer price inflation between Member States, unaffected by differences in the way it was measured. A harmonised CPI was needed because the national indices had developed over the years in many different ways, reflecting national needs and circumstances. Though they met national needs, they were demonstrably not comparable with each other" (Eurostat, 2018, p. 17).

In order to maximize comparability of the HICP among countries of the EU, there is a common framework of product categorization for all of the countries. *Classification of Individual Consumption According to Purpose* (COICOP), developed by the United Nations, is this framework. A modification of COICOP named ECOICOP is used for the HICP as well as for all other EU consumer expenditure statistics. "This classification divides the basket of goods and services into divisions (2-digit), groups (3-digit) and classes (4-digit). Further work in the EU has resulted in a refinement of COICOP to incorporate an additional level of (5-digit) sub-classes" (Eurostat, 2018, p. 27). EU member states are required by EU regulation to ensure representativeness of each ECOICOP category and inclusion of all product categories constituting at least 0.1 % of total consumption expenditure in the HICP. Weights of categories below the 0.1 % threshold are redistributed towards other categories. The following chart provides a detailed example of the ECOICOP classification.

Figure 2.1: Classification structure for various levels of aggregation of the HICP



Source: Eurostat (2018)

Eurostat (2018) describes the process of calculation of the HICP as follows.

> "Calculating the HICP is a two-stage process. The first stage consists of computing price indices for each of the elementary aggregates (i.e. the elementary price indices) within the classification structure. The second stage involves arranging these elementary price indices into a number of higher-level indices up to and including the all-items HICP. Aggregating through to the all-items level is accomplished by taking weighted averages of the lower-level indices, where the individual weight for each of these indices is equal to the expenditure share for the corresponding good or service" (Eurostat, 2018, p. 34).

One of the features distinguishing the HICP from some national CPI indices is the exclusion of imputed rent from the index calculation. The HICP puts emphasis on pure price change and covers only realized monetary transactions. The methods of measuring expenditure by owner-occupiers when purchasing housing vary among countries, so its inclusion would worsen the comparability of the HICP. This important distinction may explain the possible difference between an inflation rate calculated with the use of the HICP and an inflation rate calculated with the use of a national CPI.

This difference may be noteworthy at times. According to the Czech National Bank (2021), the difference between domestic CPI and HICP inflation in the Czech Republic throughout 2021 was enlarged up to 1 percentage point, with HICP inflation rate being the one with lower figures. This gap is a consequence of imputed rent being present in the national CPI calculation (with quite a significant weight of more than 10 % of the consumer basket), while being absent from the HICP calculation for the Czech Republic (Czech National Bank, 2021). The growth of imputed rent in the Czech Republic exceeded 10 % year-over-year in September 2021 owing to high growth in property prices, construction prices and construction materials prices in particular (Czech National Bank, 2021).

Other notable consumption segments missing from the HICP are games of chance, some types of insurance and goods and services associated with the shadow economy. Household non-consumption expenditure such as capital acquisition and interest payments is excluded from the HICP or from any consumer price index for that matter.

Prices of products such as petrol, fish or fruits and vegetables tend to be quite volatile in the short run. It is recommended that such prices are collected multiple times a month instead of just once in order to smooth short-term price fluctuations (Eurostat, 2018). Non-discriminatory discounts on individual goods or services known to the consumers in advance are included in the HICP calculation, meaning that they get deducted from the full prices. Discounts available only to a select group of the population are generally disregarded for purposes of the index. There is an important time distinction between goods and services; prices of goods are entered into the HICP calculation immediately in the month corresponding to their observation and prices of services are entered into the calculation only after consumption of the service has started. For some services such as hotel accommodation or airplane tickets, this time difference may be substantial.

Sources for weighting of the HICP sub-indices are primarily national accounts data and household budget surveys. National accounts data are preferred for higher level weights because they are already structured around the ECOICOP classification and are available yearly, while household budget surveys are frequently used for assignment of weights to product groups at lower levels of aggregation in the consumption basket, such as elementary product groups and elementary aggregates (Eurostat, 2018). However, household budget surveys are associated with several issues. Their results may not be statistically reliable if they have a small sample size. They exclude non-residents, which may cause under-reporting of consumption of some products. They have also been unreliable indicators of household expenditure on alcohol and tobacco because both of these goods tend to be under-reported as well (Eurostat, 2018).

Market research or scanner data are sometimes additionally used for weighting purposes of product groups at lower levels of aggregation. A potential problem regarding the use of these methods is that the collected data may include sales to businesses (Eurostat, 2018). These sales are not covered by any consumer price index and their inclusion in the index calculation is therefore not desirable.

National accounts data and household budget surveys can both also be used for estimation of geographical weights of outlets used in the representative consumption sample. These geographical weights may be assigned to each region of a given country, or they may only express the distinction between urban and rural consumption or, perhaps more accurately, urban and rural shopping patterns.

These shopping patterns may not reflect the distribution of the population because of the geographical phenomenon of outshopping. “Outshopping is the behavior of consumers who make retail purchases outside of the community in which they live. Outshopping is most prevalent in rural communities and represents a serious threat to rural retailers” (Broekemier & Burkink, 2004, p. 63). Home (2002) studied shopping behavior of Finnish rural consumers and found that only about a third of rural households shopped for groceries primarily in rural stores. Those engaging in outshopping tended to have a larger household size, likely due to the presence of children, and higher income and also tended to be younger, more educated and more likely to drive a car for shopping (Home, 2002).

In order to remain representative, the weighting of outlet types used by consumers should reflect outshopping as well as the noticeable partial shift of consumers’ shopping patterns to online channels. 74 % of internet users in the EU purchased goods or services online at least once during 2021 (Eurostat, 2022b). Van Droogenbroeck and Van Hove (2017) found in a survey of customers of a Belgian supermarket chain that personal and especially household characteristics similar to those of rural outshoppers, namely higher level of education, presence of young children in the household and a high relative number of full-time working adults in the household, likewise increase the probability of adoption of online grocery shopping. It may be the case that the COVID-19 pandemic has incentivized online grocery shopping and online shopping in general. For example, Switzerland experienced a steep growth of e-commerce and online grocery shopping during the lockdown in spring 2020, with 56 % of online food retailers reporting greater than 20% year-over-year growth of online orders (Zumstein & Oswald, 2020).

The consumption basket is at least partially annually reweighted because EU regulation requires member states to update the weights of ECOICOP 2-digit to 5-digit levels of aggregation (divisions, groups, classes and sub-classes) every year. These are the four highest levels of aggregation. On the other hand, lower level weights do not need to be updated as often. EU regulation requires the data on which weights for elementary product groups and elementary aggregates are based not to be more than 7 years old. This reflects the lower frequency of conducting household budget surveys in some member states. More specifically, the consumption basket is resampled each December. Apart from adjustment of weights of product groups, there is a possibility of adding newly significant products, product groups or consumption segments to the basket and removing obsolete ones. Eurostat (2018) summarizes the process of resampling as follows.

> “Resampling involves an overlap period (i.e. the December of each year) in which the old and new samples are both priced. This forms the basis for annual chain-linking. In the HICP, prices for both the old and new samples are simultaneously collected in December of each year. The index starting with the following January is calculated using only the new price sample” (Eurostat, 2018, p. 56).

Product offers (individual price observations) in the sample can be replaced at any time, even outside of the regular reweighting process, and with immediate effect if the products in question cease to be available in outlets or no longer represent a significant part of household consumption. EU regulation requires that the replacement product belongs to the same consumption segment and is selected with a focus on similarity of use compared to the product being replaced and not on similarity of prices of these products.

With potentially quite frequent replacements of products in the consumption basket consequently arises the issue of quality change of the consumption basket over time. Quality can be described as a set of traits of a product, especially those traits that influence consumers' decision about purchasing the product. Because the aim of the HICP is to measure pure price change, quality adjustment needs to take place to prevent bias of the index, as was discussed previously. A quality adjustment is made by modifying either the current price of a product or the price from a previous period in order to express the change in quality of the product in monetary terms. Dalén (1998) offers a theoretical framework of such quality adjustment. In this framework, quality adjustment is interpreted as a *g factor*. At first, total price change ($T_{t-1,t}$) is defined as

$$T_{t-1,t} = \frac{p_{B,t}}{p_{A,t-1}}$$

in which $p_{A,t-1}$ is the price of product A in period $t-1$ and $p_{B,t}$ is the price of product B in period t. The replacement of product A by product B may occur due to reasons specified previously. Due to the replacement of one product by another, the prices are not directly comparable to each other in a sense of measuring pure price change. In order to obtain pure price change ($I_{t-1,t}$), a decomposition

$$T_{t-1,t} = I_{t-1,t} \cdot g$$

of total price change is used. The symbol g, also called *g factor*, represents the quality adjustment. If $g = 1$, there is no quality change present in the sample. The rearranged form

$$g = \frac{T_{t-1,t}}{I_{t-1,t}} = \frac{p_{B,t}}{p_{A,t-1}} \cdot \frac{1}{I_{t-1,t}}$$

of the previous equation can also express the *g factor*. This framework is applicable at various levels of aggregation of the index.

Even more broadly at the level of the whole consumption basket, the total price change is equivalent to a hypothetical HICP without any quality adjustments. Such an index is named the standard reference index (SRI) and its calculation consists only of averaging the price changes within the consumer basket and weighting the consumption segments and ECOICOP levels in the same way as with HICP weighting. In a similar manner, pure price change is equivalent to the HICP itself, as the main stated aim of the HICP is to be a measure of pure price change. The *g factor*, which is perceived as quality adjustment of individual products or product groups at a lower level of aggregation, can at the general level be made equivalent to the implicit quality index (IQI). As per Eurostat (2018),

$$IQI = \frac{SRI}{HICP}$$

is thus a more general analogy to the previous equation expressing the *g factor*. Incorrect application of quality adjustment may result in a biased index series. Inadequate quality adjustment may result in an upward bias of the index and hence exaggerated inflation figures during periods of greater quality growth of some consumption basket elements, as shown for example in studies conducted by Shapiro and Wilcox (1996) and Lebow and Rudd (2003). Downward bias may be introduced into the index series if fashion variation is falsely perceived as quality improvement (Eurostat, 2018).

2.3 PPI

The PPI is a price index measuring the average change over time in the selling prices received by domestic producers for their output. It is the oldest continuous statistical series published by the United States Bureau of Labor Statistics with the data for its calculation being collected since 1890 (Samuelson & Nordhaus, 2009). The PPI is a less prominent measure of inflation than the CPI. The PPI is a Laspeyres index published monthly, same as the CPI, and it is also concerned with a representative fixed basket of goods and services, although with respect to production.

These are the major similarities of the indices. However, there are also some significant differences, which are emphasized by the US BLS (2021). The PPI is concerned with domestic output, which is also a domain of the GDP deflator but not the CPI, and focuses on price change from the perspective of the seller, as opposed to focus on the cost of living and the buyer's perspective in the case of the CPI. The PPI excludes imports and includes exports; it is vice versa for the CPI. Sales to businesses as inputs to production (including capital investment) and government purchases are also within the scope of the PPI; neither of those are included in the CPI calculation.

Taxes on consumption are handled differently by the two indices, as is succinctly described on the US BLS website.

> "The price collected for an item included in the PPI is the revenue received by its producer. Sales and excise taxes are not included in the price because they do not represent revenue to the producer. The price collected for an item included in the CPI is the out-of-pocket expenditure by a consumer for the item. Sales and excise taxes are included in the price because they are necessary expenditures by the consumer for the item" (U.S. Bureau of Labor Statistics, 2021).

While Barro (2007) claims that the PPI does not cover services and primarily includes goods that are raw materials and semi-finished products, at least in the United States, the index has been expanding its coverage into the tertiary sector of the economy since the mid-1980s (U.S. Bureau of Labor Statistics, 2021). Nevertheless, the focus of the PPI on raw materials and semi-finished products is still warranted because it measures prices at the wholesale stage of production. In the United States, it was called the wholesale price index (WPI) before its name changed to the current one in 1978. The PPI is to a great extent influenced by business-to-business transactions because consumer demand for raw materials and unfinished products is likely negligible.

Blanchard et al. (2010) in their macroeconomic textbook with a European perspective name the following as major industries covered by the PPI: manufacturing, mining, agriculture, fishing, forestry and electric utility industries. This list is almost identical to the one specified by Betsock and Gerduk (1993) for the US context. PPI sub-indices for food and energy have historically exhibited greater short-term volatility than PPI sub-indices for goods other than food and energy (U.S. Bureau of Labor Statistics, 2021). Food and energy prices may be excluded from the PPI calculation in order to create the core PPI and clearly show PPI trends undisturbed by the short-term volatility of these problematic sectors. The US BLS calculates a version of the PPI without food, energy and also trade services, whose prices are also volatile in the short term (U.S. Bureau of Labor Statistics, 2021).

The coverage of services in general appears to be more varied among countries and regions compared to the coverage of industries. While Betsock and Gerduk (1993) specify that the US PPI was at that time expanding coverage into the transportation, communication and services sectors, the Eurostat introduced its services producer price index (SPPI) only in 2006 (Eurostat, 2022c) and Richardson (2009) categorizes the UK SPPI merely as experimental and notes that it is not classified as a National Statistic in the UK.

Similarly to sub-indices of the CPI, there are many sub-indices of the PPI for product groups on different levels of aggregation, which are assigned weights in the index calculation in order to reflect the share of different product groups in the PPI basket. According to the US BLS (2021), there are about 10000 PPI sub-indices for individual products and groups of products that the Bureau releases each month. The PPI product groups in the EU are sorted and aggregated by a framework analogous to the ECOICOP classification for consumer goods and services − the NACE (abbreviation of the French *Nomenclature générale des activités économiques dans les Communautés européennes*). The English name for the NACE is *Statistical Classification of Economic Activities in the European Community*.

Eurostat (2012) specifies that the NACE framework includes classes (4-digit level), groups (3-digit level), divisions (2-digit level), sections (1-letter level) and the overall index as the main levels of aggregation. This is a very similar hierarchy to the ECOICOP; the only notable difference is that the 1-letter level of sections is missing in the ECOICOP. There are also some minor aggregation levels below the 4-digit classes and at the lowest level, there are price observations for individual products. The current version of the NACE has been in use since 2009 after a major revision of the framework and serves as the basis of price surveys (Eurostat, 2012). The sections are denoted by letters A−U and are the final level of aggregation below the overall index; they often have quite a broad scope as shown for example by section C, which indicates manufacturing. Some of the NACE sections, specifically sections H, I, J, L, M and N, focus on services and therefore are within the scope of the SPPI before the final PPI aggregation (Eurostat, 2021).

Eurostat (2012) outlines the process of calculation of the PPI in the following way.

> “Mainly, PPIs are calculated in three stages, depending on their aggregation level. The stepwise compilation procedure implies the aggregation of the lower-level indices to

obtain the higher levels ones [sic], up to the overall index. This technique is consistent in aggregation because it should grant (apart from rounding issues) the same result as if the total index had been compiled in one step. The first step in compiling indices concerns price relatives. Each price relative is the quotient of the ratio between the current monthly price (numerator) and the base price (denominator). In a second step price relatives are aggregated to obtain the elementary price index, also known in the PPI context as the elementary product index. In the third and final step the elementary product indices are aggregated as weighted averages (typically as a Laspeyres-type index) to provide a set of synthetic indices up to the overall index" (Eurostat, 2012, p. 122).

The PPI is intended to be a measure of pure price change and therefore quality adjustment is also a part of the index calculation process.

The sources for the construction of the PPI in the vast majority of EU member states are exclusively statistical surveys (Eurostat, 2012). The surveys are conducted using questionnaires sent to reporting units. Data collection in almost all countries of the EU takes place monthly. Collection of producer price data separated by markets into domestic and non-domestic is required of all EU member states; the total PPI is then obtained by weighting the domestic and non-domestic sub-indices. Additionally, countries in the euro area have to provide a further breakdown of the non-domestic market sub-index between the euro area markets and the non-euro area markets. However, there are substantial differences in the depth of the collected data; some countries provide data only for the NACE sections (1-letter level of aggregation) and the largest countries generally provide data detailed down to the NACE classes (4-digit level) (Eurostat, 2012).

Contrary to the CPI, which gets resampled every year in the EU, there is a noticeable variability in approach to resampling of the PPI basket, which is described by Eurostat (2012).

> "A total of 12 countries reported that they update the sample of observation units every five years; three of these indicated that minor changes are made more frequently. Three countries indicate that they continuously update the sample - this may be monthly or annually. The remainder generally update every year or every two years" (Eurostat, 2012, p. 42).

Research into bias of the PPI seems less noteworthy than research into bias of the CPI. As the PPI is a Laspeyres index, it may suffer from bias in a similar fashion as discussed in the CPI section. Woods (2008) specifically connects the fixed basket (Laspeyres) approach to PPI construction with substitution bias, new item bias and quality change bias within the PPI. Chopova et al. (2012) analyze the possible presence of non-response bias in PPI data in the United States and find that only 3.7 % of industries (relatively narrow product groups) show statistically significant signs of non-response bias. The authors conclude that PPI indices as a whole do not suffer from non-response bias. A brief comparison of development and general trends in the PPI and in the CPI and a discussion of a possible causal relationship between the two indices is left to Chapter 3.

2.4 GDP deflator

Although the GDP deflator is not the main topic of this thesis unlike the CPI and the PPI, it is nevertheless important to briefly mention some of its distinctions. The GDP deflator is defined as the ratio of nominal GDP to real GDP and GDP deflator inflation can be expressed as the difference between the growth rate of nominal GDP and the growth rate of real GDP (Blanchard, Amighini, & Giavazzi, 2010).

The GDP deflator is a Paasche index unlike both the CPI and the PPI. As mentioned in the beginning of this chapter, a Paasche index is a price index, which uses quantities of goods and services purchased (or rather domestically produced in the case of the GDP deflator) in the current period as weights. Using the same notation style as previously with the CPI formula, the value of a Paasche index in period *t* is equal to the ratio of the cost of the current period basket (corresponding to period *t*) to the cost of the current period basket in the base period (*b*). All values for quantity (q_i, $i \in \{1, \dots, n\}$) of products in a Paasche index are taken from the current period (*t*). The cost of the basket in each time period (*b*, *t*) is equal to the sum of products of price (p_i, $i \in \{1, \dots, n\}$) and quantity of each good or service included in the index calculation. Then, the formula

$$P_t^{Paasche} = \frac{p_{1,t}q_{1,t} + p_{2,t}q_{2,t} + \cdots + p_{n,t}q_{n,t}}{p_{1,b}q_{1,t} + p_{2,b}q_{2,t} + \cdots + p_{n,b}q_{n,t}} = \frac{\sum_{i=1}^{n} p_{i,t}q_{i,t}}{\sum_{i=1}^{n} p_{i,b}q_{i,t}}$$

is the expression of the construction of a Paasche index in mathematical terms.

It is also previously mentioned in this chapter that a Paasche index understates inflation. More specifically, by projecting the current consumption basket backward into some past period, a Paasche index effectively assumes that the present substitution effect already occurred in the past (Schultze, 2003). In spite of this, "it is at least conceptually possible that a change in the Paasche index could exceed a change in the Laspeyres index" (Schultze, 2003, p. 6).

The scope of the GDP deflator differs from that of the CPI in two major ways. Some part of domestic production of the economy, which is covered by the GDP deflator, may not be sold to consumers but instead to firms, the government or it may be exported, leaving it outside the scope of the CPI, while some part of domestic consumption, which is covered by the CPI, is in the form of imports, which are outside the scope of the GDP deflator (Blanchard, Amighini, & Giavazzi, 2010). Both the CPI and the GDP deflator have been tracked for decades in developed economies and their time series may provide an interesting comparison.

Garín et al. (2018) find using United States quarterly data from the period 1947−2016 that the GDP deflator shows on average a lower level of inflation than the CPI. In particular, the average inflation rate during the period is 0.8 % per quarter and about 3.2 % annualized for the GDP deflator and 0.9 % per quarter and about 3.6 % annualized for the CPI (Garín, Lester, & Sims, 2018). This situation regarding these measures of inflation may be expected due to the already described differences between the biases of Laspeyres and Paasche indices. Herr and Kazandziska (2010) find that even in a special case of the stagnating and often deflationary economy of 1990s and 2000s Japan, the GDP deflator has lower and more quickly decreasing index numbers than the CPI, which in this situation means that the GDP deflator expresses higher deflation instead of lower inflation. Although the GDP deflator of Japan decreased by more than 12 % during the period 1994–2008, the Japanese CPI decreased only slightly during the same period (Herr & Kazandziska, 2010). The GDP deflator also generally exhibits lower volatility than the CPI (Garín, Lester, & Sims, 2018).

3 Literature review

Before summarizing the papers focusing on the relationship of the PPI and the CPI, it may be helpful to mention some major differences in the development of the two indices. Vermeulen et al. (2007) find using data from six euro area countries that producer prices are more flexible (meaning that price changes have a higher frequency) than consumer prices in all countries included in the paper and that price changes, both upwards and downwards, tend to be larger for consumer prices than for producer prices. The former of these findings is supported by Polák and Novotný (2020), who note that measured by standard deviation, the PPI in the EU in the period from January 2001 to August 2020 was about three times more volatile than the CPI. The authors highlight the importance of commodities to high volatility of the PPI and further explain its volatility by its strong ties to economic sectors most sensitive to changes in the business cycle and economic shocks, such as industry. However, the development of both indices in the EU has been fairly similar since the Great Recession (Polák & Novotný, 2020).

New methods of exploration of causal relationships among economic variables came into existence in the second half of the 20th century. Such econometric tools include the Granger causality test or VAR models. One of the earlier papers on the relationship between the PPI and the CPI by Colclough and Lange (1982) responds to a preceding paper by Silver and Wallace (1980) whose analysis, as the authors claim, did not properly utilize the econometric methods used to determine causality. Colclough and Lange (1982) start their analysis by describing a possible theoretical mechanism of causality from consumer prices to producer prices, which states that it is the demand for final goods and services that generates the opportunity costs of resources and intermediate materials used in production, and also acknowledge that this direction of causality is contrary to commonplace assertion. The data used in the analysis are monthly percent changes in the US CPI and the US PPI (called WPI at the time) in the period 1945–1979. Both the Sims test and the Granger test employed by the authors support bidirectional causality between the two indices.

A key paper focusing on the topic of price index causality is that of Clark (1995). The author notes that conventional economic theory suggests the existence of a pass-through effect in the production chain going from producer prices to consumer prices, with firms setting the price of their output as a markup over the cost of production. Given the markup, a change in cost will cause the price to change as well according to this theory. The author also acknowledges theoretical weaknesses of this connection between producer prices and consumer prices by highlighting the differences between the PPI and the CPI, which are described in more detail in Chapter 2, and gives increased domestic production costs combined with decreased import prices as an example of a situation where the PPI and the CPI may develop separately.

The empirical part of the paper starts with a graphical analysis of historical movements of three PPI sub-indices (for crude materials, for intermediate goods, for finished goods) compared to the CPI. The data is quarterly from the period 1958–1994 in the US. The surface-level graphical comparison shows only weak and occasional links from any of the PPI sub-indices to the CPI; however, as the author notes, the large volatility of the PPI might make any pass-through effect of producer price changes to consumer prices difficult to discern from the charts. A simple regression is then used to show that lagged changes in PPI inflation are statistically significant in explaining variation in core CPI inflation as well as core goods CPI inflation.

The final level of the empirical analysis includes two VAR forecasting models with three additional variables apart from the indices. These variables are real GDP growth, the 3-month Treasury bill rate, and growth in the average hourly earnings of manufacturing workers. One of the models includes both the PPI and the CPI, the other includes only the CPI. The explained variables are either the core CPI or the core goods CPI. The data is split into two groups. Values from the period 1959–1976 are used for model estimation; values from the period 1977–1994 are used for model estimation only if they precede the forecasted quarter, otherwise they constitute out-of-sample data used to compare the predictions forecasted by the models to actual levels of inflation.

In a given time period, performance of the models is evaluated by average one-year-ahead absolute forecast errors. For the whole period 1977–1994, the average absolute error is indeed lower for the model with producer prices. However, in some shorter periods within, the model without the PPI performs better. This applies regardless of whether the explained variable is the core CPI or the core goods CPI. The author concludes that PPI changes sometimes help predict CPI changes but fail to do so systematically.

Caporale et al. (2002) study the relationship between producer prices and consumer prices in the countries of the G7. The authors note that the PPI is widely used as a leading indicator for the CPI because of supply-side developments, in which the retail sector adds value with a lag to existing production. The authors stress that an omitted relevant variable in bivariate VAR models may result in invalid inference about the causality structure of the included variables. The first model of the empirical analysis is a bivariate VAR including the logarithms of the PPI and the CPI. The data is quarterly from the period 1976–1999 from each country of the G7 for both models. The results of this model are mixed, with Canada exhibiting no causality at all, France and Germany having unidirectional causality running from the PPI to the CPI and Italy, Japan, the UK and the US exhibiting bidirectional causality.

The second VAR model consists of five variables in order to capture the transmission mechanism of monetary policy and deal with the omitted variable issue, which makes it more suitable than the bivariate model according to the authors. The additional variables are the money supply, real GDP and the three-month interest rate. The results of the five-variate model support unidirectional causality from the PPI to the CPI in all countries. Therefore, Caporale et al. (2002) support the predominant view that producer prices lead consumer prices, as the five-variate model is preferred.

This thesis focuses in part on the relationship between the PPI and the CPI in the Czech Republic. Notable papers regarding this topic using Czech data are Holub (2000) and Khan et al. (2018). The data in Holub (2000) come from the period 1993–1997. This period can be characterized by its high inflation as the country was transitioning away from a command economy towards a market economy. Indeed, Holub (2000) shows graphically that the Czech CPI inflation rate in the period in question peaked above 20 % in 1993 and spent a great part of the period settled at just under 10 %, staying above 5 % at all times.

Holub (2000) uses various methods, starting with correlation analysis. The time series are first-differenced in order to achieve stationarity. Already at this level, a conclusion can be made that the PPI is certainly not an indicator for prices of housing and food. The next level of analysis are Granger causality tests; the PPI Granger-causes the CPI without housing and food prices with a lag of about 3 months. The final level of analysis is a multivariate consumer price inflation model with variables for the money supply, wages and the exchange rate besides producer prices. Again, the lagged values of the PPI are statistically significant for explaining the CPI with housing and food prices removed but not the total CPI. A similar multivariate producer price model also shows causality from the CPI to the PPI. Although it appears that the overall Granger causality is therefore unidirectional from the CPI to the PPI, the explanatory power of the PPI towards relevant CPI segments is also quite strong.

The other paper by Khan et al. (2018) focuses on ten Central and Eastern European countries including the Czech Republic. The authors find Granger causality from the PPI to the CPI in Latvia, Lithuania, Romania, Slovakia and Slovenia and Granger causality in the other direction only in Hungary; there is no relationship in either direction in the Czech Republic. Overall, there is mixed evidence regarding the relationship between the PPI and the CPI in the Czech Republic considering the conclusions of these two papers.

Other papers on this topic use a variety of methods and study quite a diverse group of countries, both developed and developing. Ghazali et al. (2008) analyze the relationship in Malaysia using Granger causality tests on data from the period from January 1986 to April 2007 and find unidirectional causality from the PPI to the CPI. Fan et al. (2009) employ Granger causality tests and the distributed lag model on data from the period from January 2001 to July 2008 in China and find that the CPI Granger-causes the PPI with a lag of 1–3 months, which they attribute to stronger influence of demand-side factors compared to supply-side factors in the Chinese economy.

Two of the papers are concerned with Mexico. In the first one, Sidaoui et al. (2009) use the vector error correction (VEC) model on data from the period from June 2000 to June 2009 to find evidence of long-run unidirectional Granger causality from the PPI to the CPI but no short-run Granger causality, with both of these results then being corroborated by out-of-sample forecasts for the period from June 2003 to June 2009. Specifically, the PPI helps to forecast CPI inflation for horizons beyond 8 months. In the second paper, Tiwari et al. (2014) follow up on this research through the means of the wavelet transform method and find bidirectional causality dependent on the time horizon; the PPI leads the CPI in the long run (8–32-month period), which corroborates Sidaoui et al. (2009), but the CPI leads the PPI in the short run.

The same method is used by Tiwari et al. (2013) on data from the period 1991–2011 in Romania and no stable causal relationship is observed; the authors instead suggest the presence of cyclical effects because the causality appears to fluctuate depending on the particular sub-sample. Woo et al. (2019) study the relationship between the PPI and the CPI in the UK, France and Germany in the period 1997–2013 using the TVEC model, which can be a way to examine nonlinear Granger causality, and find long-run bidirectional causality between the indices. Brazil is also represented in the research by da Rocha Lima Filho (2019), who uses the VAR and BVAR models to find unidirectional causality from the PPI to the CPI.

To summarize, regional and temporal differences among economies may entail different economic phenomena, which may be behind some of the conflicting conclusions in the literature. Unidirectional Granger causality running from producer prices to consumer prices seems to be the most frequent conclusion in the literature, but there is definitely not a consensus. The relationship between the PPI and the CPI may appear quite unstable at times. As shown by Caporale et al. (2002), an omitted variable may greatly influence the results of a model estimation and may even lead to an erroneous conclusion regarding causality within the model.

3.1 Commodity prices and inflation

There is also notable research focusing on potential causality from commodity prices towards consumer prices. Commodity prices are of course an essential component of the PPI calculation. Blomberg and Harris (1995) use VAR models with various US commodity price indices to explore this relationship. The authors conclude that the nature of the relationship changed in time. Most commodity price indices had explanatory power towards consumer price inflation in the 1970s and early 1980s but then lost it. The models surprisingly estimate a negative relationship between most commodity price indices and core CPI inflation in the period 1987–1994.

The authors explain this phenomenon in the following way: "Commodities have declined in importance, both as a share of final output and as a source of exogenous shocks to the economy. Some commodity price signals may also have been offset by countervailing changes in monetary policy" (Blomberg & Harris, 1995, p. 33). Their findings are corroborated by Furlong and Ingenito (1996), who note that if the past relationship between commodity prices and inflation valid in the 1970s and early 1980s was used to predict the CPI in the then recent years, such a prediction of US consumer price inflation would be much higher than the reported inflation in the period. Furlong and Ingenito (1996) also suggest that a change in the composition of price-affecting commodity shocks took place, particularly an increase in explanatory power of oil price shocks towards inflation.

Further developments may have brought about a new approach to the link between commodity prices and consumer prices. Bernanke (2008) emphasizes this link in light of the at that time peaking 2000s commodities boom, which was accompanied by higher levels of inflation and in which both the daily closing price and the intraday price of oil reached their historical highs of USD 145.29 per barrel and USD 147.2 per barrel respectively (Sornette, Woodard, & Zhou, 2009). Both of these events happened within just weeks of Bernanke's speech.

The relationship is revisited for instance by Gelos and Ustyugova (2012), who compare data from tens of both advanced and developing countries in the period 2001–2011 including structural characteristics of their economies and their monetary and exchange rate regimes. The authors find that developing countries experience stronger effects of commodity price shocks on inflation than countries with an advanced economy. The difference between the effects in advanced and developing economies is more pronounced for food price shocks rather than fuel price shocks and the fuel and food price pass-through effects are larger in more fuel-intensive economies and in economies with a higher share of food in the CPI basket (Gelos & Ustyugova, 2012).

4 Data and methodology

4.1 Data

We examine the relationship between the CPI and the PPI in both the Czech Republic and the euro area. The euro area expanded multiple times during the period covered by our euro area dataset. It grew from 12 countries at the start of the dataset to 19 countries at the end of the dataset with the accession of Slovenia in 2007, Cyprus and Malta in 2008, Slovakia in 2009, Estonia in 2011, Latvia in 2014 and Lithuania in 2015 (European Central Bank, n.d.). However, these changes presumably do not need addressing because all the newly accessed countries together constitute only about 2.5 % of euro area HICP country weights as of 2019 (Eurostat, 2022a). It is therefore unlikely that the accession of any of these countries into the euro area would influence the common euro area statistics in a significant manner.

For both the Czech Republic and the euro area, we estimate a bivariate and a multivariate model for the purpose of Granger causality testing. The bivariate models include only industrial PPI inflation and either CPI (for the Czech Republic) or HICP (for the euro area) inflation. It is important to note that in the euro area models, the variable HICP is labelled as CPI in order to unify the naming of the variables across all models.

In particular, the variables are entered in the form of year-over-year percentage change of the index numbers similarly to Clark (1995) and Fan et al. (2009), as opposed to the approach of Caporale et al. (2002), who use a logarithmic transformation of the index numbers instead. As these indices are monthly statistics, the bivariate models include monthly observations. The model for the Czech Republic includes values from the period 2006–2021 and the model for the euro area includes values from the period 2001–2020.

The multivariate models include three additional variables in both of our regions of interest. The model for the Czech Republic includes, apart from PPI and CPI inflation, the real GDP growth rate, the growth rate of the M2 monetary aggregate and the growth rate of average gross nominal wages. The model for the euro area includes, apart from PPI and HICP inflation, the real GDP growth rate, the growth rate of the M3 monetary aggregate and the growth rate of the LCI. All of these variables are entered in the form of a year-over-year percentage change.

As mentioned in Chapter 3, Caporale et al. (2002) show the importance of involving monetary policy developments while testing for Granger causality between price indices. This is the reason for inclusion of the growth rates of the monetary aggregates in our multivariate models. The use of the other two variables added into the multivariate models is motivated by Clark (1995), who uses the real GDP growth rate as a variable signifying the overall state of the economy and includes growth in the average hourly earnings of manufacturing workers as a variable covering the labor market. While we include the real GDP growth rate as well, labor market developments in our multivariate models are encompassed by the growth rate of average gross nominal wages and the growth rate of the LCI, respectively. “These variables are included because they may be related in different ways to both producer and consumer price inflation. Including these variables helps identify pass-through effects that might otherwise be obscured” (Clark, 1995, p. 32).

However, some of the additional variables are only available quarterly at most. The monthly time series from the bivariate model therefore need to be transformed into quarterly time series, which is done by using a standard procedure in which the observations in the months at the end of quarters represent the values for the whole quarter. In particular, observations for March, June, September and December are taken as observations for Q1, Q2, Q3 and Q4, respectively. Observations for other months are disregarded, which considerably shortens the length of the dataset. The data is sourced from the Czech Statistical Office, the CNB, Eurostat and the ECB. A supplementary table located in the Appendix summarizes the data used in the two models for either region.

4.2 Methodology

The framework of approaching our research question is Granger causality, first introduced by Granger (1969). The main goal of our analysis is to determine whether lagged values of PPI inflation help predict CPI inflation, lagged values of CPI inflation help predict PPI inflation (either option would be an instance of unidirectional Granger causality), both variables help predict each other (bidirectional Granger causality) or whether there is no relationship between PPI inflation and CPI inflation and thus no Granger causality. All models are estimated using OLS.

Firstly, we focus on whether the time series are stationary. Stationarity is an important trait of time series data. Non-stationary time series suffer from a high persistence of shocks and their use in econometric analysis can lead to spurious regressions (Brooks, 2008). A stationary stochastic process has a joint probability distribution that is stable over time. A less strict condition is named covariance stationarity and is fulfilled when a process has a constant mean, a constant finite variance and a constant autocovariance structure, which means that for any $t, h \geq 1$ in a stochastic process $\{x_t: t = 1, 2, \dots\}$, $Cov(x_t, x_{t+h})$ depends only on h and not on t (Brooks, 2008; Wooldridge, 2012). Covariance stationarity suffices for the purpose of standard econometric analysis (Brooks, 2008).

We use the Dickey-Fuller test (DF) and its version the augmented Dickey-Fuller test (ADF). These are common unit root tests. A unit root process is a type of a non-stationary process. The mechanism of the DF test may be explained following Enders (2009). For a time series $\{y_t: t = 1, 2, \dots\}$, we are interested in whether $a_1 = 1$ in the AR(1) model of $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$. A subtraction of y_{t-1} from each side of the AR(1) regression leads to the equation

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$

in which $\gamma = a_1 - 1$. Consequently, testing the hypothesis $a_1 = 1$ is equivalent to testing the hypothesis $\gamma = 0$ (Enders, 2009).

The DF test has the following hypotheses.

H_0: $\gamma = 0$ $(a_1 = 1)$ – $\{y_t\}$ contains a unit root (and therefore is non-stationary).

H_1: $\gamma < 0$ $(a_1 < 1)$ – $\{y_t\}$ does not contain a unit root and is a I(0) process.

The option that $\gamma > 0$ ($a_1 > 1$) is not considered because such an autoregressive process would be explosive. A process containing one unit root is also labelled I(1) (integrated of order 1) and a process not containing a unit root is labelled I(0) (integrated of order 0). Although it is possible for a stochastic process not to contain a unit root and still be non-stationary, rejection of the null hypothesis strongly suggests stationarity in common financial and macroeconomic time series.

We use a type of the DF test with an intercept (a_0, also called drift term) and without a deterministic linear time trend. We use the version with a drift because it is common to leave the intercept unspecified under the null hypothesis (Wooldridge, 2012). We also use the ADF test, which is commonly used instead of the DF test in order to address potential serial correlation in disturbances of the DF regression equation. The ADF test leads to the equation

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

in which p lags of the dependent variable are added into the regression. The lags of Δy_t ensure that the errors ε_t do not exhibit serial correlation (Brooks, 2008). Nothing regarding interpretation of the ADF test changes from the DF test; the ADF test is still conducted on γ and still has the same hypotheses and critical values as the DF test. A new issue is the choice of optimal lag length p. A typical recommendation is to decide based on the frequency of the data, which means including 4 lags of the dependent variable for quarterly data and 12 lags for monthly data (Brooks, 2008; Wooldridge, 2012), however, "there are no hard rules to follow in any case" (Wooldridge, 2012, p. 642). Another option is to select the number of lags that minimizes the value of an information criterion such as the Akaike information criterion (AIC) or Schwarz information criterion (SIC) (Brooks, 2008).

"The most important criticism that has been levelled at unit root tests is that their power is low if the process is stationary but with a root close to the non-stationary boundary" (Brooks, 2008, p. 330). This may very well be the case with our data. Brooks (2008) suggests to tackle this issue by complementing the unit root test such as the ADF test with a stationarity test such as the KPSS test. The KPSS test reverses the testing approach of the unit root tests and has a null hypothesis of I(0) (and thus stationarity in common financial and macroeconomic data) against the alternative hypothesis of non-stationarity. Lag length is also a part of the KPSS test specification, but our software (R) selects the lag length automatically and somewhat arbitrarily for two variants – short and long.

An advantage of including a stationarity test in the analysis is that the data will appear stationary by default if there is little information in the sample, as there is not enough evidence to reject the null hypothesis (Brooks, 2008). But if there is not enough evidence to reject the null hypothesis of a unit root test, the data will instead appear non-stationary by default. If such a conflicting combination of results were to occur, it would show us that the information included in our data is not enough to determine if the data are I(0) or I(1), possibly due to small sample size. However, if we used only a unit root test, we would have, perhaps inaccurately, concluded that our data was non-stationary.

An important testing step for I(1) time series is a cointegration test. “The notion of cointegration applies when two series are I(1), but a linear combination of them is I(0); in this case, the regression of one on the other is not spurious, but instead tells us something about the long-run relationship between them” (Wooldridge, 2012, p. 632). We use the Engle-Granger two-step method as a cointegration test and we follow Fan et al. (2009) in doing so. For two I(1) time series, it requires two (contemporaneous) regressions of one variable on the other. For our purpose,

$$PPI_t = \alpha_1 + \beta_1 CPI_t + \varepsilon_{1t}$$

is the first regression and

$$CPI_t = \alpha_2 + \beta_2 PPI_t + \varepsilon_{2t}$$

is the second regression. The residuals from these two regressions are then tested for a unit root by the DF test. The variant without a drift is used as the expected value of the residuals is 0. If the null hypothesis of the DF test is rejected, the residuals are stationary and this means that the two variables are cointegrated.

As mentioned previously, the framework of our research question is Granger causality because we are interested in the direction of the possible relationship between the PPI and the CPI. Having two variables *x* and *y* with their time series $\{x_t\}$ and $\{y_t\}$, *x* Granger-causes *y* if

$$E(y_t|I_{t-1}) \neq E(y_t|J_{t-1}),$$

where I_{t-1} contains information on past values of *x* and *y* and J_{t-1} contains only information on past values of *y* (Wooldridge, 2012). The inequality means that past values of *x* help predict y_t after controlling for past values of *y*.

In practice, testing for Granger causality means estimating some regression model. Using the notation of Holub (2000), the regression equation for testing if *x* Granger-causes *y* is

$$y_t = c + \sum_{j=1}^{k} \alpha_j \, y_{t-j} + \sum_{j=1}^{k} \beta_j \, x_{t-j} + \varepsilon_t$$

for $j \in \{1, \dots, k\}$. Determining if *x* Granger-causes *y* is done by testing the joint significance of coefficients β_j by an F-test. The Granger causality test has the following hypotheses.

H_0: $\beta_j = 0 \; \forall j$ – *x* does not Granger-cause *y*.

H_1: β_j are jointly significant – *x* Granger-causes *y*.

In a multivariate setting, we rely on an expanded definition of Granger causality named conditional Granger causality. It still describes a relationship between two time series, just conditional on a third time series or a set of time series. The meaning of the test does not change and nor do the hypotheses; the data is merely added into the models in the form of lagged values as one or more additional independent variables. Our multivariate models are examples of testing for conditional Granger causality; we are interested in Granger causality from the PPI to the CPI or vice versa conditional on lagged values of real GDP growth, growth of a monetary aggregate and a labor market variable, respectively.

Choice of lag length (also called lag order) k is also a part of the testing procedure. It does not appear to be particularly strict, similarly to the choice of lag length for the ADF test. Fan et al. (2009) merely employ the Granger causality test for every lag from 1 to 18 and report the results separately by the number of lags with an emphasis on those lags where the null hypothesis is rejected. Other options are using the convention that 4 lags are included for quarterly data and 12 lags for monthly data or minimization of an information criterion.

For example, Clark (1995) selects the lag length using the AIC and Holub (2000) uses both the AIC and the SIC. Clark (1995) remarks that the SIC yields models with much shorter lag lengths than the AIC. At least for vector autoregression (VAR) models, which are rather related to Granger causality and especially conditional Granger causality, this observation is corroborated by Lütkepohl (2011), who claims that the AIC selects the largest lag length, the SIC selects the smallest lag length and a third criterion, the Hannan-Quinn information criterion (HQIC), selects a lag length in between the two. This happens in a situation when the criteria do not agree with each other. In general, they may all select the same lag length (Lütkepohl, 2011).

If heteroskedasticity was present, it would be possible to use heteroskedasticity robust standard errors in our estimation. However, we do not need to test for heteroskedasticity because the goal of the estimation of the models is testing for Granger causality. This in practice means an F-test of joint significance of some coefficients, as is established in this chapter. For that, we do not necessarily need to have precise estimates of the value of each coefficient in our regression, as we do not need to forecast these concrete values. Similarly, we also do not use the ARIMA model in our analysis, which is used mainly in such forecasting. ARIMA models also cannot be estimated by OLS because they contain unobservable independent variables in the form of lagged errors. Regarding serial correlation, it appears that it cannot be involved under the null hypothesis of the Granger causality test if the *correct* number of lags is used (Wooldridge, 2012).

5 Empirical study

In this chapter, we summarize the results of the estimation of our models. The samples from the Czech Republic are not from the same time period as the samples from the euro area, although they considerably overlap. The reason is primarily difference in data availability between the Czech Republic and the euro area for some variables in the multivariate models. We also acknowledge the limitations of our approach regarding the issue of causality at the end of this chapter.

Observations for 2021 in the euro area were not considered because the increase in PPI inflation during 2021 was rather extreme in context of the rest of the sample and values of PPI inflation, especially at the end of the year, would be strong outliers. The same could be stated about inflation figures of both the PPI and the CPI in both the Czech Republic and the euro area in 2022. 2022 has exhibited levels of inflation that are unprecedented in the 21st century. Data for 2022 is thus not considered at all, even though it is the most recent available data.

The source of the tables in this chapter is own calculation of the author. In tables with missing p-values, statistical significance is expressed by one asterisk at the 10 % level, two asterisks at the 5 % level and three asterisks at the 1 % level. Regarding the descriptive statistics, some are rounded to one decimal point and some are not depending on whether the growth rates, which are the form in which the variables are presented, come from own calculation from the index numbers or directly from the source. This minor distinction should not have any effect on the results of the conducted statistical tests. Much of the commentary about the bivariate models applies to the multivariate models as well.

5.1 Results

5.1.1 Bivariate models

The bivariate model for the Czech Republic includes 192 monthly observations from the period 2006–2021. The variable names are *PPI* and *CPI* and they are in the form described in the previous chapter. Following are some descriptive statistics of the variables.

Table 5.1: Descriptive statistics – Czech Republic

Variable	Minimum	Mean	Maximum	St. dev.
PPI (% y-o-y)	-5.44	1.46	13.51	3.35
CPI (% y-o-y)	-0.11	2.29	7.64	1.64

The bivariate model for the euro area includes 240 monthly observations from the period 2001–2020. Again, variables *PPI* and *CPI* are in the form described in the previous chapter and the following table contains descriptive statistics of the variables.

Table 5.2: Descriptive statistics – euro area

Variable	Minimum	Mean	Maximum	St. dev.
PPI (% y-o-y)	-7.30	1.20	7.30	2.88
CPI (% y-o-y)	-0.60	1.63	4.10	0.98

The mean of *CPI* is higher than the mean of *PPI* by 0.83 percentage points in the Czech Republic and by 0.43 percentage points in the euro area. This may be the case due to a lower deflationary spike of the PPI in 2009 following the Great Recession, due to PPI deflation in a period of low CPI inflation around 2014–2016 and, lastly, due to the samples not being from the exact same time periods. The standard deviation of *PPI* is in both cases between 2–3 times higher than the standard deviation of *CPI*, which is an expected ratio based on Polák and Novotný (2020).

The variables do not exhibit seasonality to an extent that would warrant seasonal differencing or an addition of seasonal dummy variables. The variables in the Czech model seem seasonal, but after removing the observations from 2021, when inflation rose considerably, the supposed seasonality largely disappears.

Unit root and stationarity testing is conducted by the DF test, the ADF test and the KPSS test. All of these tests are described in detail in Chapter 4. With the unit root tests, which are the former two, we use a variant with an intercept (drift term) and without a time trend. Brooks (2008) opines that it is more common for time series in finance and economics to look like a random walk with drift rather than a time series with a deterministic time trend. Analogous to that, we use the variant of the KPSS test that tests for level stationarity, not trend stationarity. The results of the unit root tests are as follows.

Table 5.3: DF test – Czech Republic

Variable	Test statistic	Specification
PPI	0.3308	Intercept, no trend
CPI	-0.904	Intercept, no trend

Table 5.4: ADF test – Czech Republic

Variable	Test statistic	Lags	Lags chosen by	Specification
PPI	-1.4493	12	Convention (12), AIC	Intercept, no trend
PPI	-1.632	1	SIC	Intercept, no trend
CPI	-1.2895	12	Convention (12), AIC, SIC	Intercept, no trend

Table 5.5: DF test – euro area

Variable	Test statistic	Specification
PPI	-2.0747	Intercept, no trend
CPI	-1.6262	Intercept, no trend

Table 5.6: ADF test – euro area

Variable	Test statistic	Lags	Lags chosen by	Specification
PPI	-2.0572	12	Convention (12), AIC, SIC	Intercept, no trend
CPI	-1.9239	12	Convention (12), AIC	Intercept, no trend
CPI	-2.6061*	2	SIC	Intercept, no trend

Critical values for the DF test and the ADF test with an intercept and no time trend are -2.57 for 10 % significance level, -2.86 for 5 % significance level and -3.43 for 1 % significance level (Wooldridge, 2012).

For both variables *PPI* and *CPI*, the null hypothesis of a unit root mostly cannot be rejected by the results of the DF and ADF tests, apart from euro area *CPI*, where there is a weak rejection of the null hypothesis with a particular lag length chosen by the SIC. However, for reasons discussed previously, using also a test with the null hypothesis of stationarity, such as the KPSS test, enhances the testing.

Table 5.7: KPSS test – Czech Republic

Variable	Test statistic	Lags	p-value
PPI	0.2811	4	>0.1
PPI	0.13563	14	>0.1
CPI	0.40558	4	0.07
CPI	0.18773	14	>0.1

Table 5.8: KPSS test – euro area

Variable	Test statistic	Lags	p-value
PPI	0.59348	4	0.02
PPI	0.29916	14	>0.1
CPI	1.7369	4	<0.01
CPI	0.75159	14	<0.01

The null hypothesis of stationarity of *PPI* in the Czech Republic is not rejected, which is a result conflicting with the unit root tests, which do not reject the null hypothesis of a unit root. For *CPI* in the Czech Republic, the evidence points slightly towards non-stationarity with a weak rejection of stationarity for one of the lag lengths of the KPSS test combined with a non-rejection of the unit root in unit root tests. The same can be stated about *PPI* in the euro area, only the rejection of stationarity is stronger for one of the lag lengths, but the test with a longer lag length still does not reject stationarity. Testing of *CPI* in the euro area yields conflicting results, this time with a strong rejection of stationarity in the KPSS test combined with a weak rejection of an unit root for one of the lag lengths in the ADF test.

Overall, the results are far from decisive, although they seem to favour non-stationarity of the Czech *CPI*, euro area *PPI* and euro area *CPI*. Only for Czech *PPI* is the testing entirely inconclusive. We are aware of the fact that non-stationarity of time series may lead to invalid statistical inference. A cautious approach would lead one to consider all of our time series non-stationary. A standard solution would then be to take first differences of our time series, which would be I(0) if our time series are I(1).

However, first differencing is also not without issues because it may cause loss of information in the data. Sims et al. (1990) indicate that first differencing of data "whenever it appears likely that the data are integrated is in many cases unnecessary" (Sims, Stock, & Watson, 1990, p. 136). "When the preliminary tests suggest a particular nonstationary form for the model but at a marginal p-value of, say, .10 or .15, one could consider tests of the hypotheses of interest both under the integrated and nonintegrated maintained hypotheses" (Sims, Stock,

& Watson, 1990, p. 137). *Integrated* in the context of these quotes means I(1). This suggestion may also be relevant to our data because of the inconclusive test results.

Ultimately, we base our treatment of the possible non-stationarity of PPI and CPI inflation on the approach of Fan et al. (2009) to precisely these variables in China. This approach entails testing the two variables for cointegration using the Engle-Granger two-step method. Assuming they are both I(1), which may even be likely, if they are cointegrated, we can use them in the Granger causality test regardless of the unit root. We follow the manner in which Fan et al. (2009) use this method, as outlined in Chapter 4. The results are as follows.

Table 5.9: Cointegration test – Engle-Granger two-step method – Czech Republic

Regression	Test statistic	Specification
PPI~CPI	-1.5843	No intercept, no trend
CPI~PPI	-2.3952**	No intercept, no trend

Table 5.10: Cointegration test – Engle-Granger two-step method – euro area

Regression	Test statistic	Specification
PPI~CPI	-2.967***	No intercept, no trend
CPI~PPI	-2.6124***	No intercept, no trend

Critical values for the DF test with no intercept and no time trend (which is the final part of the Engle-Granger two-step method) are -1.61 for 10 % significance level, -1.94 for 5 % significance level and -2.59 for 1 % significance level (Fan, He, & Hu, 2009).

In the model for the Czech Republic, one of the test statistics is very close to the value of the 10 % significance level and the other indicates significance at 5 % level. It can be assumed that *PPI* and *CPI* are cointegrated (if they are both I(1)), as there is enough evidence to reject the null hypothesis of no cointegration.

In the model for the euro area, both of the test statistics indicate significance at 1 % level. It can be assumed that *PPI* and *CPI* are cointegrated (if they are both I(1)) with more certainty than in the model for the Czech Republic; there is substantial evidence to reject the null hypothesis of no cointegration.

After these preliminary tests, we proceed with the Granger causality test. The results are reported in similar manner to Fan et al. (2009). As mentioned, the Granger causality test explores both directions and in our case, it can be characterized by the equations

$$CPI_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i CPI_{t-i} + \sum_{i=1}^{k} \beta_i PPI_{t-i} + \varepsilon_t$$

$$PPI_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i PPI_{t-i} + \sum_{i=1}^{k} \beta_i CPI_{t-i} + \varepsilon_t$$

in which $i \in \{1, \dots, k\}$ and *k* is the lag order.

The Granger causality test characterized by the first equation has the following hypotheses.

H_0: $\beta_i = 0\ \forall i$ – *PPI* does not Granger-cause *CPI*.

H_1: β_i are jointly significant – *PPI* Granger-causes *CPI*.

The Granger causality test characterized by the second equation has the following hypotheses.

H_0: $\beta_i = 0\ \forall i$ – *CPI* does not Granger-cause *PPI*.

H_1: β_i are jointly significant – *CPI* Granger-causes *PPI*.

The results are reported for each lag from 1 up to 18, inspired by Fan et al. (2009). Additionally, there is a note for lag lengths selected by the information criteria (AIC, SIC, HQIC).

Table 5.11: Granger causality test – Czech Republic

Lag order (*k*)	Null hypothesis	F-statistic	p-value	Notes
1	I	19.163	<0.01	
1	II	5.6062	0.02	
2	I	9.4812	<0.01	SIC, HQIC
2	II	0.6939	0.50	SIC, HQIC
3	I	5.521	<0.01	
3	II	0.7023	0.55	
4	I	4.7229	<0.01	
4	II	0.4875	0.74	
5	I	3.7843	<0.01	
5	II	0.5743	0.72	
6	I	3.2636	<0.01	
6	II	0.6194	0.71	
7	I	2.8818	<0.01	
7	II	0.6795	0.69	
8	I	2.4042	0.02	
8	II	0.7717	0.63	
9	I	1.9425	0.05	
9	II	0.6014	0.79	
10	I	1.8203	0.06	
10	II	0.534	0.86	
11	I	1.5594	0.12	
11	II	1.0888	0.37	
12	I	1.3707	0.19	
12	II	1.0124	0.44	
13	I	1.2811	0.23	
13	II	0.6805	0.78	
14	I	1.1562	0.31	AIC
14	II	0.6389	0.83	AIC
15	I	0.9656	0.49	
15	II	0.6555	0.82	
16	I	0.9173	0.55	
16	II	0.7545	0.73	
17	I	0.9325	0.54	
17	II	1.0135	0.45	
18	I	0.7906	0.71	
18	II	0.9409	0.53	

Null hypothesis I: *PPI* does not Granger-cause *CPI*. Null hypothesis II: *CPI* does not Granger-cause *PPI*.

Table 5.12: Granger causality test – euro area

Lag order (*k*)	Null hypothesis	F-statistic	p-value	Notes
1	I	4.2032	0.04	
1	II	0.9829	0.32	
2	I	18.774	<0.01	SIC
2	II	0.0653	0.94	SIC
3	I	9.7918	<0.01	
3	II	0.0228	1.00	
4	I	8.6019	<0.01	HQIC
4	II	0.2522	0.91	HQIC
5	I	7.2765	<0.01	
5	II	0.4176	0.84	
6	I	5.7934	<0.01	
6	II	0.717	0.64	
7	I	5.0747	<0.01	
7	II	0.6898	0.68	
8	I	5.0023	<0.01	
8	II	1.0695	0.39	
9	I	4.3746	<0.01	
9	II	0.8623	0.56	
10	I	4.0729	<0.01	
10	II	1.0236	0.42	
11	I	3.6563	<0.01	
11	II	1.1202	0.35	
12	I	3.4797	<0.01	
12	II	0.847	0.60	
13	I	2.1374	0.01	
13	II	0.7684	0.69	
14	I	2.4634	<0.01	AIC
14	II	0.6226	0.84	AIC
15	I	2.4019	<0.01	
15	II	0.6104	0.86	
16	I	2.2203	<0.01	
16	II	0.4269	0.97	
17	I	2.0524	0.01	
17	II	0.3839	0.99	
18	I	1.9338	0.02	
18	II	0.4459	0.98	

Null hypothesis I: *PPI* does not Granger-cause *CPI*. Null hypothesis II: *CPI* does not Granger-cause *PPI*.

5.1.2 Multivariate models

The multivariate model for the Czech Republic includes 64 quarterly observations from the period 2006–2021. The multivariate model for the euro area includes 80 quarterly observations from the period 2001–2020. Both of the multivariate models include five variables and three variables are the same in both models. The variable names in the Czech model are

PPI, *CPI*, *RealGDP*, *M2* and *NomW*. The variable names in the euro area model are *PPI*, *CPI*, *RealGDP*, *M3* and *LCI*. The variables are described in Chapter 4 and details about all of the variables are also listed in the Appendix. The descriptive statistics for PPI and CPI inflation should be broadly similar to the statistics of these variables in the bivariate model because they come from the same data. The data is only transformed from monthly to quarterly in order to align with the quarterly variables.

Table 5.13: Descriptive statistics – Czech Republic

Variable	Minimum	Mean	Maximum	St. dev.
PPI (% y-o-y)	-5.44	1.52	13.19	3.43
CPI (% y-o-y)	0.00	2.31	7.13	1.65
RealGDP (% y-o-y)	-10.80	2.10	9.10	3.67
M2 (% y-o-y)	0.10	7.88	17.00	4.41
NomW (% y-o-y)	-2.00	4.68	10.10	2.72

Table 5.14: Descriptive statistics – euro area

Variable	Minimum	Mean	Maximum	St. dev.
PPI (% y-o-y)	-6.80	1.17	6.60	2.83
CPI (% y-o-y)	-0.30	1.63	3.90	0.99
RealGDP (% y-o-y)	-14.20	0.86	3.70	2.60
M3 (% y-o-y)	-0.30	5.44	12.10	3.05
LCI (% y-o-y)	0.80	2.40	4.00	0.81

The statistics for the other variables are less relevant for our purposes because we still are interested in the relationship between the PPI and the CPI, only this time in the framework of conditional Granger causality. Nevertheless, as they are contained in the regression equations of the conditional variant of the Granger causality test, they are also subject to unit root and stationarity testing. Treatment of seasonality need not be considered because it was also unnecessary for monthly data and some of the additional variables are already seasonally adjusted.

Table 5.15: DF test – Czech Republic

Variable	Test statistic	Specification
PPI	-0.736	Intercept, no trend
CPI	-1.376	Intercept, no trend
RealGDP	-2.9484**	Intercept, no trend
M2	-1.6051	Intercept, no trend
NomW	-3.6246***	Intercept, no trend

Table 5.16: ADF test – Czech Republic

Variable	Test statistic	Lags	Lags chosen by	Specification
PPI	-1.5187	4	Convention (4), AIC, SIC	Intercept, no trend
CPI	-1.1713	4	Convention (4), AIC, SIC	Intercept, no trend
RealGDP	-2.2806	4	Convention (4), AIC, SIC	Intercept, no trend
M2	-2.5037	4	Convention (4)	Intercept, no trend
M2	-2.3487	2	AIC, SIC	Intercept, no trend
NomW	-1.6771	4	Convention (4)	Intercept, no trend
NomW	-1.899	3	AIC, SIC	Intercept, no trend

Table 5.17: DF test – euro area

Variable	Test statistic	Specification
PPI	-2.7485*	Intercept, no trend
CPI	-1.7946	Intercept, no trend
RealGDP	-3.4053**	Intercept, no trend
M3	-0.5448	Intercept, no trend
LCI	-2.3436	Intercept, no trend

Table 5.18: ADF test – euro area

Variable	Test statistic	Lags	Lags chosen by	Specification
PPI	-2.4346	4	Convention (4), AIC	Intercept, no trend
PPI	-4.3211***	1	SIC	Intercept, no trend
CPI	-1.9456	4	Convention (4), AIC	Intercept, no trend
CPI	-2.6753*	1	SIC	Intercept, no trend
RealGDP	-2.9225**	4	Convention (4)	Intercept, no trend
RealGDP	-3.0651**	1	AIC, SIC	Intercept, no trend
M3	-1.6399	4	Convention (4), AIC, SIC	Intercept, no trend
LCI	-2.2418	4	Convention (4)	Intercept, no trend
LCI	-2.3368	1	AIC, SIC	Intercept, no trend

Critical values for the DF test and the ADF test with an intercept and no time trend are -2.57 for 10 % significance level, -2.86 for 5 % significance level and -3.43 for 1 % significance level (Wooldridge, 2012).

The presence of a unit root cannot be rejected for any variables in the Czech model considering that the ADF test is generally more reliable than the DF test. It is also not rejected for *M3* and *LCI* in the euro area. Testing euro area *PPI* and *CPI* produces conflicting results depending on the lag length. However, there is strong evidence that *RealGDP* in the euro area is stationary. We then proceed with the KPSS test.

Table 5.19: KPSS test – Czech Republic

Variable	Test statistic	Lags	p-value
PPI	0.15692	3	>0.1
PPI	0.12993	10	>0.1
CPI	0.21361	3	>0.1
CPI	0.15282	10	>0.1
RealGDP	0.11633	3	>0.1
RealGDP	0.10062	10	>0.1
M2	0.26816	3	>0.1
M2	0.14472	10	>0.1
NomW	0.30686	3	>0.1
NomW	0.14628	10	>0.1

Table 5.20: KPSS test – euro area

Variable	Test statistic	Lags	p-value
PPI	0.32378	3	>0.1
PPI	0.28582	11	>0.1
CPI	0.87287	3	<0.01
CPI	0.52927	11	0.04
RealGDP	0.18178	3	>0.1
RealGDP	0.17276	11	>0.1
M3	0.45703	3	0.05
M3	0.21943	11	>0.1
LCI	0.81508	3	<0.01
LCI	0.38352	11	0.08

The null hypothesis of stationarity cannot be rejected for any variables in the Czech model. Combined with the results of the unit root tests, this means that testing was not able to determine whether any of the variables in the Czech model are I(0) or I(1). This is not surprising considering the performance of the tests in the bivariate case and the fact that the multivariate model for the Czech Republic has the lowest number of observations.

Regarding results of the tests of the euro area variables, there is overwhelming evidence of stationarity of *RealGDP* and relatively convincing evidence of non-stationarity of *LCI*. For *M3*, the evidence tilts towards non-stationarity, but is not entirely conclusive. The evidence for *PPI* and *CPI* is mixed. *PPI* may be the close to stationarity, but the unit root is not rejected for one of the lag lengths of the ADF test. Almost inversely, *CPI* may be close to non-stationarity but the ADF test mildly rejects the unit root for one of the lag lengths.

Overall, the reasoning from the bivariate models section applies to multivariate models as well because there is a similar degree of uncertainty about the results. We enhance the testing by again using the Engle-Granger two step method, which shows whether *PPI* and *CPI* are cointegrated if both of them are I(1).

Table 5.21: Cointegration test – Engle-Granger two-step method – Czech Republic

Regression	Test statistic	Specification
PPI~CPI	-2.052**	No intercept, no trend
CPI~PPI	-2.5555**	No intercept, no trend

Table 5.22: Cointegration test – Engle-Granger two-step method – euro area

Regression	Test statistic	Specification
PPI~CPI	-3.0711***	No intercept, no trend
CPI~PPI	-2.1869**	No intercept, no trend

Critical values for the DF test with no intercept and no time trend (which is the final part of the Engle-Granger two-step method) are -1.61 for 10 % significance level, -1.94 for 5 % significance level and -2.59 for 1 % significance level (Fan, He, & Hu, 2009).

All of the test statistics indicate significance at least at 5 % level. In both models, we can conclude that provided *PPI* and *CPI* are both I(1), they are cointegrated. As was the case with the bivariate models, we can proceed without any transformation of the variables. The regression equations for testing of conditional Granger causality are very similar to the bivariate case. The null hypothesis for each equation is the same as in the bivariate case, the only difference is the addition of *k* lags of the other independent variables into the equations.

In the model for the Czech Republic, we are testing Granger causality between *PPI* and *CPI* conditional on *RealGDP*, *M2* and *NomW*. In the model for the euro area, we are testing Granger causality between *PPI* and *CPI* conditional on *RealGDP*, *M3* and *LCI*. We have smaller sample sizes than in bivariate models because we use quarterly data and the additional variables with their lagged values also decrease the number of degrees of freedom in the regression estimations. The results are reported for each lag only up to 4 because of these limitations. A note highlighting lag lengths selected by the information criteria (AIC, SIC, HQIC) is again included in the tables.

Table 5.23: Granger causality test – Czech Republic

Lag order (*k*)	Null hypothesis	F-statistic	p-value	Notes
1	I	10.06	<0.01	SIC
1	II	3.27	0.08	SIC
2	I	7.82	<0.01	
2	II	2.44	0.1	
3	I	6.69	<0.01	
3	II	2.65	0.06	
4	I	9.08	<0.01	AIC, HQIC
4	II	3.37	0.02	AIC, HQIC

Null hypothesis I: *PPI* does not Granger-cause *CPI*. Null hypothesis II: *CPI* does not Granger-cause *PPI*.

Table 5.24: Granger causality test – euro area

Lag order (*k*)	Null hypothesis	F-statistic	p-value	Notes
1	I	1.05	0.31	SIC, HQIC
1	II	1.64	0.21	SIC, HQIC
2	I	1.05	0.36	AIC
2	II	0.95	0.39	AIC
3	I	0.65	0.59	
3	II	1.20	0.32	
4	I	0.42	0.80	
4	II	1.70	0.16	

Null hypothesis I: *PPI* does not Granger-cause *CPI*. Null hypothesis II: *CPI* does not Granger-cause *PPI*.

5.2 Discussion

Four regression models constitute the output of our analysis. For both the Czech Republic and the euro area, there is a bivariate model with data only for the PPI and the CPI and a multivariate model with the PPI, the CPI and additional economic variables. These models are used to test Granger causality between the PPI and the CPI. All previous lags are present in the regression for a particular lag order, as can be observed from the Granger causality test equations.

The results from the Czech models unambiguously suggest that the PPI is a leading indicator of the CPI in the Czech Republic. The evidence is more limited for a vice versa relationship.

In the bivariate model for the Czech Republic, we reject the null hypothesis of no Granger causality from the PPI to the CPI for lags 1–7 at 1 % significance level and for lags 8 and 9 at 5 % significance level. This is solid evidence of Granger causality in the short to medium term of 1–9 months. The null hypothesis of no Granger causality from the CPI to the PPI is rejected only for lag 1 at 5 % significance level. This result can be considered a little suspicious because for all other lags (2–18), the p-values of the F-statistics are quite high.

In the multivariate model for the Czech Republic, the results point very strongly towards Granger causality from the PPI to the CPI and partially also towards the opposite direction of the relationship, which would then take the form of bidirectional Granger causality. The rejection of the null hypothesis of no Granger causality from the PPI to the CPI is once again quite strong at 1 % significance level for all 4 quarterly lags included in the testing. The short to medium term of 1–4 quarters roughly corresponds to the monthly lags in the bivariate model for which Granger causality from the PPI to the CPI was also found. Granger causality in the other direction has less of a strong evidence. The null hypothesis for lags 1–3 is rejected at 10 % significance level and only for the lag of 4 quarters is it rejected at 5 % significance level. These results are not exactly in agreement with the results of the bivariate model, where there is Granger causality only for the shortest lag length of 1 month.

The literature is also not entirely conclusive about the relationship between the PPI and the CPI in the Czech Republic. Our findings completely differ from Khan et al. (2018), who do not find Granger causality in either direction. The findings of Holub (2000) are also different from ours because we do not eliminate prices of housing and food from our analysis and yet we find Granger causality from the PPI to the CPI even without this adjustment.

While the results of the Czech models have a relatively straightforward interpretation, at least regarding the stronger direction of Granger causality from the PPI to the CPI, the interpretation of the results of the euro area models may be more problematic.

In the bivariate model for the euro area, we observe substantial evidence of Granger causality from the PPI to the CPI, even stronger than in the Czech bivariate model. The null hypothesis of no Granger causality from the PPI to the CPI is rejected for every lag from 1 to 18 at 5 % significance level. It is even rejected at 1 % significance level for lags 2–12 and 14–16. The contrast with testing for the opposite direction of the relationship is stark; the null hypothesis of no Granger causality from the CPI to the PPI is not rejected for any of the lags from 1 to 18 and the p-values at most lag lengths are very high.

However, the multivariate model for the euro area yields vastly different results. This is surprising given the clear results of the bivariate model. No null hypothesis is rejected for any of the lags and the lowest p-value is 0.16. The conclusion of this model is that there is no evidence of any Granger causality between the PPI and the CPI.

It is difficult to explain such a discrepancy. Although the quarterly transformation of the data decreased the sample size of the PPI and the CPI three times, it should still be large enough for statistical testing. The introduction of other variables besides the price indices into the regressions may have a greater effect than expected. If we abide by Caporale et al. (2002), the multivariate model takes precedence over the bivariate model due to the presence of omitted variable bias in the bivariate model. On the other hand, monthly data may offer higher precision.

It is still unexpected nonetheless that the results of the testing changed so drastically between the bivariate and the multivariate model. This did not happen in the case of the models for the Czech Republic, where the strong Granger causality from the PPI to the CPI was preserved going from the bivariate model to the multivariate model and evidence for Granger causality in the other direction was mixed in both models.

Overall, it might still be appropriate to consider Granger causality from the PPI to the CPI in the euro area to some extent, as the null hypothesis in the bivariate model is rejected at 1 % significance level almost regardless of the lag length. The evidence that there is no Granger causality from the CPI to the PPI in the euro area is very clear. The closest paper to this topic may be Woo et al. (2019), who find long-run bidirectional causality between the indices for example in France and Germany. Such a result differs from the results of our analysis.

5.3 Limitations

The naming of Granger causality may unfortunately create some misunderstandings regarding this phenomenon. The econometric use of Granger causality is mainly connected to the forecasting ability of a variable towards a second variable when controlling for the information contained in lagged values of the second variable, rather than to some explicit causal sequence or causal mechanism (Zanetti, 2007). This notion can be succinctly summarized in a concrete example: “Evidence of Granger causality of wages on prices would, instead, *not* mean that wage changes *literally* cause price changes" (Zanetti, 2007, p. 68). Granger causality also does not imply anything about contemporaneous causality and is not applicable to the context of cross-sectional data (Wooldridge, 2012).

Correct understanding of these pieces of information is important in order to accurately interpret the results of the Granger causality test. It can even be compellingly argued that Granger causality is "somewhat of a misnomer" (Brooks, 2008, p. 298). Leamer (1985) describes Granger causality as "a specialized concept that is only incidentally related to causality as most of us use the term" (Leamer, 1985, p. 283) and recommends to use the word *precedence* instead.

6 Conclusion

The goal of this thesis is to analyze the relationship between the PPI and the CPI through the framework of Granger causality. We also demonstrate awareness of what Granger causality actually entails and what are its limitations. Data from the period 2006–2021 in the Czech Republic and 2001–2020 in the euro area are analyzed. Two regression models used for Granger causality tests are estimated for each of these regions; a bivariate model with just the two price indices themselves and a multivariate model with five variables. The three additional variables are included in order to control for the overall performance of the economy in some way. In particular, these variables track real GDP, development of a given monetary aggregate and wages. The choice of these variables is mostly inspired by previous literature on the topic in order to more extensively cover the economy as a whole in our models. There are four main models estimated in total and each of these is estimated at various lag orders.

Strong Granger causality, characterized by rejection of the null hypothesis of the Granger causality test at a low enough significance level, is discovered mainly from the PPI to the CPI. However, some models also hint at bidirectional Granger causality between the indices. These findings appear to mostly support the overall tendencies outlined in the literature. There may be a difference between the conclusion of the whole body of literature, which is mostly in accordance with our results, and the conclusion of papers most similar to our analysis in terms of their regional orientation. The reason for that is diversity among the conclusions of the papers. The answer to our research question appears not to be firmly set and instead may depend on regional characteristics of the economy, in which the relationship is scrutinized, and also on the time period.

Although there have been papers on this topic making use of more advanced methods, the Granger causality test is a standard and reliable tool, which is well equipped for issues like our topic. We have already established that the results may vary among regions and time periods. This is where our analysis might be useful; it concerns relatively recent Czech and euro area data. Our dataset even covers the beginning of the 2020s, an unstable period in terms of inflation, for both of these regions.

This topic in general may be of great interest to any institution focused on macroeconomics, as inflation is one of the most important macroeconomic variables, but particularly to central banks. For instance, if it indeed is true that the PPI Granger-causes the CPI in the short to medium term, as our analysis for the most part suggests, a CPI inflation targeting central bank might also be interested in somehow influencing the trajectory of the PPI. If this is successful, the PPI might then influence the CPI by proxy with a desired time lag of several months. The lag would most likely be dependent on regional macroeconomic specifics. It is also not surprising that some of the more prominent sources on our topic are working papers originating from central banks, particularly the Federal Reserve System.

The analysis may be extended by using more advanced methods than Granger causality. Some of them are mentioned in Chapter 3. There is a relatively large array of possible methods. The other ways to improve the analysis are inclusion of more variables in the multivariate regression models in order to further mitigate the omitted variable bias and also larger sample size in order to better preserve the degrees of freedom in the estimations. However, any research on this topic may not realistically avoid being bound to a certain region, time period and economic realities stemming from these two factors. In our opinion, it is evidently unlikely that there will ever be research on the relationship of the PPI and the CPI that is universal to all regions and time periods, as demonstrated by the often contradictory nature of the literature.

7 References

Anghelache, C., Manole, A., Diaconu, A., Gheorghe, M., & Ciocan, L. (2012). The Walsh index and the theory of the "pure" price index. *Romanian Statistical Review Supplement, 60*(1), 123-126.

Barro, R. J. (2007). *Macroeconomics: A modern approach* (1st ed.). South-Western Educational Publishing.

Bernanke, B. (2008, June 9). Outstanding issues in the analysis of inflation. *Speech at the Federal Reserve Bank of Boston's 53rd Annual Economic Conference*. Chatham, Massachusetts. Retrieved November 7, 2022, from https://www.federalreserve.gov/newsevents/speech/bernanke20080609a.htm

Betsock, T., & Gerduk, I. (1993). The problem of list prices in the producer price index: The steel mill products case. In M. Foss, M. Manser, & A. Young, *Price measurements and their uses* (National Bureau of Economic Research Studies in Income and Wealth ed., pp. 261-274). Chicago: University of Chicago Press.

Blanchard, O., Amighini, A., & Giavazzi, F. (2010). *Macroeconomics: A European perspective* (1st ed.). Pearson Education.

Blomberg, S. B., & Harris, E. S. (1995). The commodity-consumer price connection: Fact or fable? *Economic Policy Review-Federal Reserve Bank of New York, 1*(3), 21-38.

Broekemier, G. M., & Burkink, T. J. (2004). Using the internet to facilitate outshopping by rural residents: Survey and implications. *Journal of Internet Commerce, 3*(1), 63-78.

Brooks, C. (2008). *Introductory econometrics for finance* (2nd ed.). Cambridge University Press.

Caporale, G. M., Katsimi, M., & Pittis, N. (2002). Causality links between consumer and producer prices: Some empirical evidence. *Southern Economic Journal, 68*(3), 703-711.

Chopova, B., Hesley, T. E., Kelly, G., Paik, S., Sadler, A., & Slack, D. (2012). *Analyzing nonresponse bias in PPI data.* Washington, D.C.: U.S. Bureau of Labor Statistics.

Clark, T. E. (1995). Do producer prices lead consumer prices? *Economic Review-Federal Reserve Bank of Kansas City, 80*(3), 25-39.

Colclough, W. G., & Lange, M. D. (1982). Empirical evidence of causality from consumer to wholesale prices. *Journal of Econometrics, 19*, 379-384.

Czech National Bank. (2021). Differences in the measurement of consumer price inflation from the point of view of the national consumer price index (CPI) versus the EU harmonised index (HICP). *Monetary Policy Report – Autumn 2021*, pp. 27-28.

da Rocha Lima Filho, R. I. (2019). Does PPI lead CPI in Brazil? *International Journal of Production Economics, 214*, 73-79.

Dalén, J. (1998). *On the statistical objective of a Laspeyres' price index.* Statistical Office of the European Communities.

Diewert, W. E. (1992). Fisher ideal output, input and productivity indexes revisited. *Journal of Productivity Analysis, 3*(3), 211-248.

Enders, W. (2009). *Applied econometric time series* (3rd ed.). Wiley.

European Central Bank. (n.d.). *Our money*. Retrieved December 9, 2022, from https://www.ecb.europa.eu/euro/intro/html/index.en.html

Eurostat. (2008). *European price statistics – An overview.* Luxembourg: Office for Official Publications of the European Communities.

Eurostat. (2012). *Handbook on industrial producer price indices (PPI).* Luxembourg: Publications Office of the European Union.

Eurostat. (2018). *Harmonised index of consumer prices (HICP) methodological manual.* Luxembourg: Publications Office of the European Union.

Eurostat. (2021). *Glossary: Services producer price index (SPPI)*. Retrieved October 7, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Services_producer_price_index_(SPPI)

Eurostat. (2022a). *HICP - country weights*. Retrieved December 9, 2022, from Eurostat Data Browser: https://ec.europa.eu/eurostat/databrowser/view/PRC_HICP_COW__custom_4105329/default/table?lang=en

Eurostat. (2022b). *Internet purchases by individuals (2020 onwards)*. Retrieved May 5, 2022, from https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do

Eurostat. (2022c). *Services producer price index overview*. Retrieved September 16, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Services_producer_price_index_overview

Fan, G., He, L., & Hu, J. (2009). CPI vs. PPI: Which drives which? *Frontiers of Economics in China, 4*(3), 317-334.

Furlong, F., & Ingenito, R. (1996). Commodity prices and inflation. *Economic Review-Federal Reserve Bank of San Francisco*, 27-47.

Garín, J., Lester, R., & Sims, E. (2018). *Intermediate macroeconomics* (3rd ed.).

Gelos, G., & Ustyugova, Y. (2012). *Inflation responses to commodity price shocks – How and why do countries differ?* IMF Working Paper No. 12/225, International Monetary Fund.

Ghazali, M. F., Ooi, A. Y., & Muhammad, M. Z. (2008). Do producer prices cause consumer prices? Some empirical evidence. *International Journal of Business and Management, 3*(11), 78-82.

Gordon, R. J., & vanGoethem, T. (2003). A century of housing shelter prices: How big is the CPI bias? *CRIW Conference in Memory of Zvi Griliches, Hard-to-Measure Goods and Services.* Bethesda, MD.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society, 37*(3), 424-438.

Greenstein, S., & McDevitt, R. (2011). Evidence of a modest price decline in US broadband services. *Information Economics and Policy, 23*(2), 200-211.

Groshen, E. L., Moyer, B. C., Aizcorbe, A. M., Bradley, R., & Friedman, D. M. (2017). How government statistics adjust for potential biases from quality change and new goods in an age of digital technologies: A view from the trenches. *Journal of Economic Perspectives, 31*(2), 187-210.

Herr, H., & Kazandziska, M. (2010). The labour market and deflation in Japan. *International Journal of Labour Research, 2*(1), 79-98.

Holub, T. (2000). Prices of industrial outputs and inflationary developments in the Czech Republic. *Eastern European Economics, 38*(3), 6-25.

Home, N. (2002). Rural consumers' patronage behaviour in Finland. *The International Review of Retail, Distribution and Consumer Research, 12*(2), 149-164.

Khan, K., Su, C. W., Tao, R., & Lobonţ, O. R. (2018). Producer price index and consumer price index: Causality in Central and Eastern European countries. *Ekonomický časopis, 66*(4), 367-395.

Kryvtsov, O. (2016). Is there a quality bias in the Canadian CPI? Evidence from microdata. *Canadian Journal of Economics/Revue canadienne d'économique, 49*(4), 1401-1424.

Leamer, E. E. (1985). Vector autoregressions for causal inference? *Carnegie-Rochester Conference Series on Public Policy. 22*, pp. 255-304. North-Holland.

Lebow, D. E., & Rudd, J. B. (2003). Measurement error in the consumer price index: Where do we stand? *Journal of Economic Literature, 41*(1), 159-201.

Lütkepohl, H. (2011). *Vector autoregressive models.* EUI Working Paper ECO 2011/30, European University Institute, Florence.

Mankiw, N. G. (2010). *Macroeconomics* (7th ed.). New York: Worth Publishers.

Polák, P., & Novotný, F. (2020). Consumer and industrial prices in 2020 – The year of the coronavirus. *Global Economic Outlook – December 2020*. Czech National Bank.

Richardson, I. (2009). Services producer price index (experimental) – Third quarter 2008. *Economic & Labour Market Review, 3*(1), 74-76.

Samuelson, P., & Nordhaus, W. (2009). *Economics* (19th ed.). New York: McGraw-Hill.

Schultze, C. L. (2003). The consumer price index: Conceptual issues and practical suggestions. *Journal of Economic Perspectives, 17*(1), 3-22.

Shapiro, M. D., & Wilcox, D. W. (1996). Mismeasurement in the consumer price index: An evaluation. *NBER Macroeconomics Annual, 11*, 93-142.

Sidaoui, J., Capistrán, C., Chiquiar, D., & Ramos-Francia, M. (2009). *A note on the predictive content of PPI over CPI inflation: The case of Mexico.* Banco de México Working Paper No. 2009-14, Banco de México.

Silver, J. L., & Wallace, T. D. (1980). The lag relationship between wholesale and consumer prices: An application of the Hatanaka-Wallace procedure. *Journal of Econometrics, 12*(3), 375-387.

Sims, C. A., Stock, J. H., & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica: Journal of the Econometric Society, 58*(1), 113-144.

Sornette, D., Woodard, R., & Zhou, W. X. (2009). The 2006–2008 oil bubble: Evidence of speculation, and prediction. *Physica A: Statistical Mechanics and its Applications, 388*(8), 1571-1576.

Tiwari, A. K., Mutascu, M., & Andries, A. M. (2013). Decomposing time-frequency relationship between producer price and consumer price indices in Romania through wavelet analysis. *Economic Modelling, 31*, 151-159.

Tiwari, A. K., Suresh, K. G., Arouri, M., & Teulon, F. (2014). Causality between consumer price and producer price: Evidence from Mexico. *Economic Modelling, 36*, 432-440.

U.S. Bureau of Labor Statistics. (2021). *Producer price indexes frequently asked questions (FAQs)*. Retrieved September 11, 2022, from https://www.bls.gov/ppi/faqs/questions-and-answers.htm

U.S. Bureau of Labor Statistics. (2022). *Consumer price index frequently asked questions*. Retrieved May 14, 2022, from https://www.bls.gov/cpi/questions-and-answers.htm

Van Droogenbroeck, E., & Van Hove, L. (2017). Adoption of online grocery shopping: Personal or household characteristics? *Journal of Internet Commerce, 16*(3), 255-286.

Vermeulen, P., Dias, D., Dossche, M., Gautier, E., Hernando, I., Sabbatini, R., & Stahl, H. (2007). *Price setting in the euro area: Some stylised facts from individual producer price data.* NBB Working Paper No. 111, National Bank of Belgium, Brussels.

Woo, K. Y., Lee, S. K., & Ng, C. Y. (2019). An investigation into the dynamic relationship between CPI and PPI: Evidence from the UK, France and Germany. *The Singapore Economic Review, 64*(5), 1081-1100.

Woods, J. (2008). Measuring the quality of the producer price index – An update. *Economic & Labour Market Review, 2*(8), 45-48.

Wooldridge, J. M. (2012). *Introductory econometrics: A modern approach* (5th ed.). Cengage Learning.

Zanetti, A. (2007). Do wages lead inflation? Swiss evidence. *Swiss Journal of Economics and Statistics, 143*(1), 67-92.

Zumstein, D., & Oswald, C. (2020). *Swiss online retailer survey 2020: The sustainable growth of e-commerce and challenges in times of crisis.* Zurich University of Applied Sciences, Institute of Marketing Management. Winterthur: ZHAW School of Management and Law.

8 Appendix

Table 8.1: Detailed data description

Czech Republic – bivariate model					
Variable	**Description**	**Unit**	**Time period**	**Frequency**	**Source**
PPI	Annual rate of change of the industrial PPI	%	2006–2021	M	CZSO Public database
CPI	Annual rate of change of the CPI	%	2006–2021	M	CZSO Public database
Czech Republic – multivariate model					
Variable	**Description**	**Unit**	**Time period**	**Frequency**	**Source**
PPI	Annual rate of change of the industrial PPI	%	2006–2021	Q	CZSO Public database
CPI	Annual rate of change of the CPI	%	2006–2021	Q	CZSO Public database
RealGDP	Annual growth rate of real GDP (seasonally adjusted)	%	2006–2021	Q	CZSO Public database
M2	Annual growth rate of the M2 monetary aggregate	%	2006–2021	Q	CNB Time series database – ARAD
NomW	Annual growth rate of average gross monthly wages and salaries (full-time equivalent, nominal)	%	2006–2021	Q	CZSO Public database
Euro area – bivariate model					
Variable	**Description**	**Unit**	**Time period**	**Frequency**	**Source**
PPI	Annual rate of change of the industrial PPI	%	2001–2020	M	Eurostat
HICP	Annual rate of change of the HICP	%	2001–2020	M	Eurostat
Euro area – multivariate model					
Variable	**Description**	**Unit**	**Time period**	**Frequency**	**Source**
PPI	Annual rate of change of the industrial PPI	%	2001–2020	Q	Eurostat
HICP	Annual rate of change of the HICP	%	2001–2020	Q	Eurostat
RealGDP	Annual growth rate of real GDP (seasonally adjusted)	%	2001–2020	Q	Eurostat
M3	Annual growth rate of the M3 monetary aggregate	%	2001–2020	Q	ECB Statistical Data Warehouse
LCI	Annual growth rate of the nominal LCI (seasonally adjusted)	%	2001–2020	Q	Eurostat

The following figures display monthly data.

Figure 8.1: Development of PPI inflation in the Czech Republic

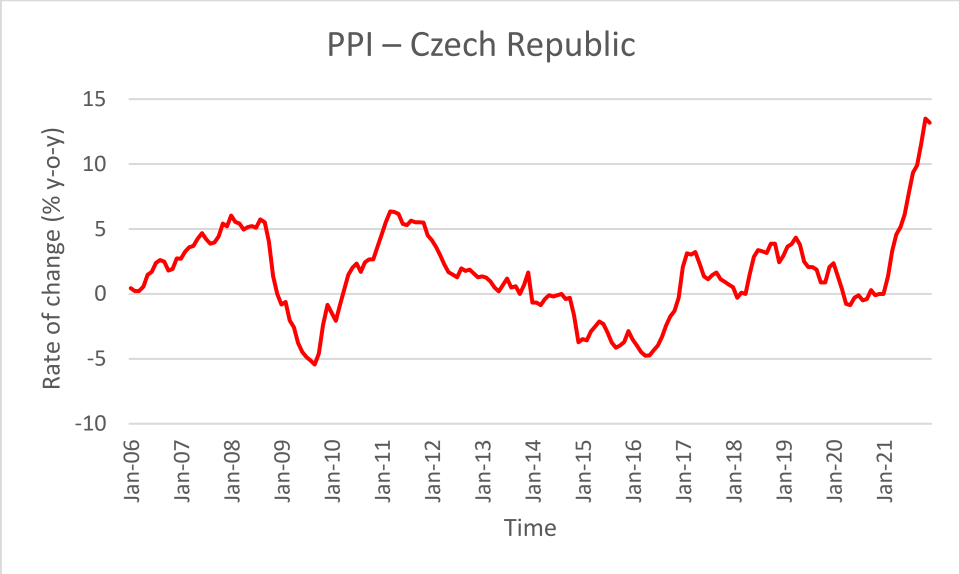


Figure 8.2: Development of CPI inflation in the Czech Republic

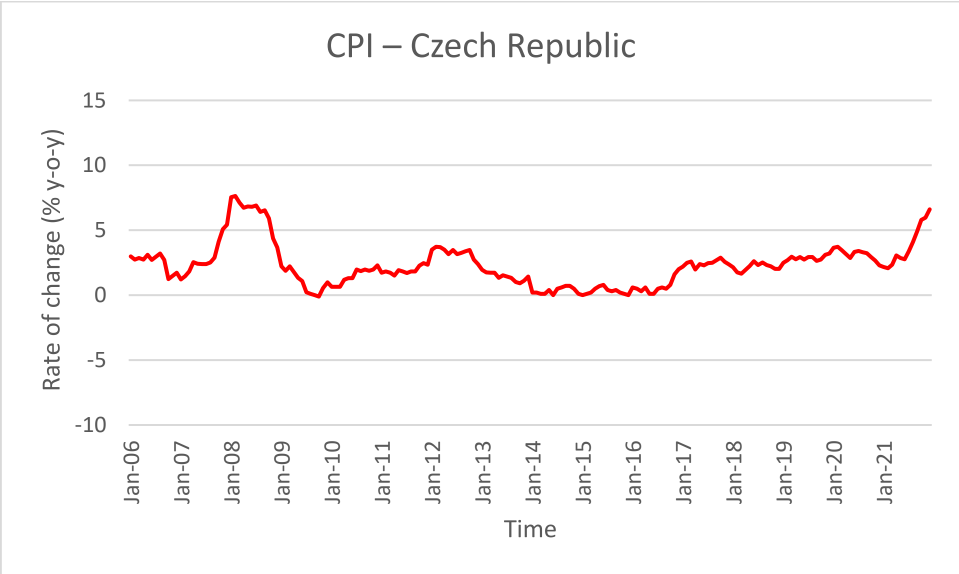


Figure 8.3: Development of PPI inflation in the euro area

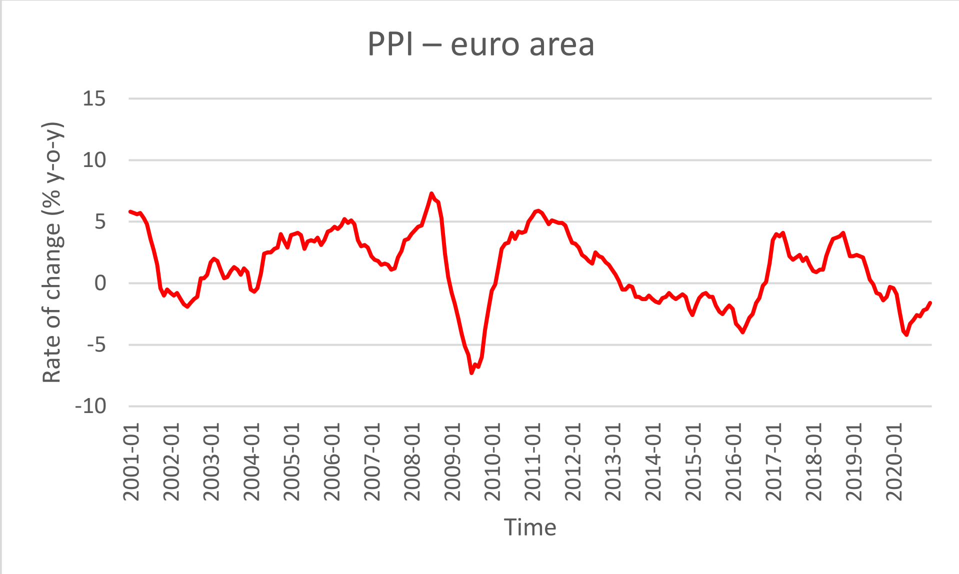


Figure 8.4: Development of CPI inflation in the euro area

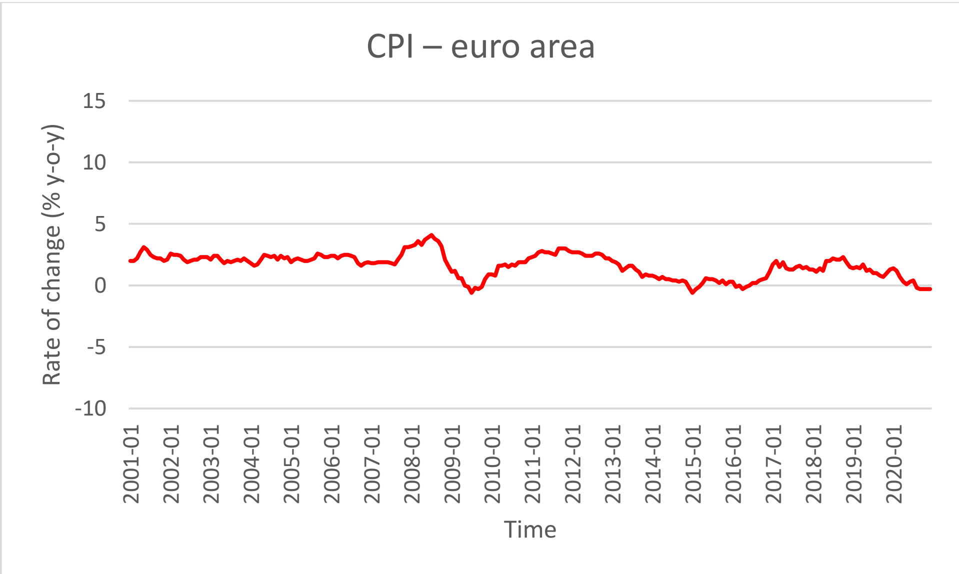
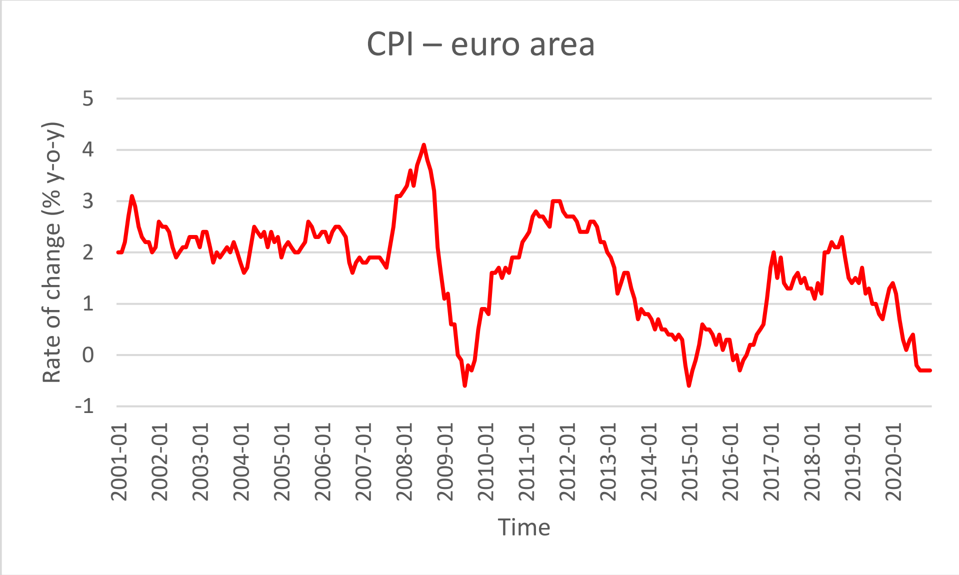


Figure 8.5: Development of CPI inflation in the euro area in more detail



The following figures display quarterly data.

Figure 8.6: Development of variable *RealGDP* in the Czech Republic

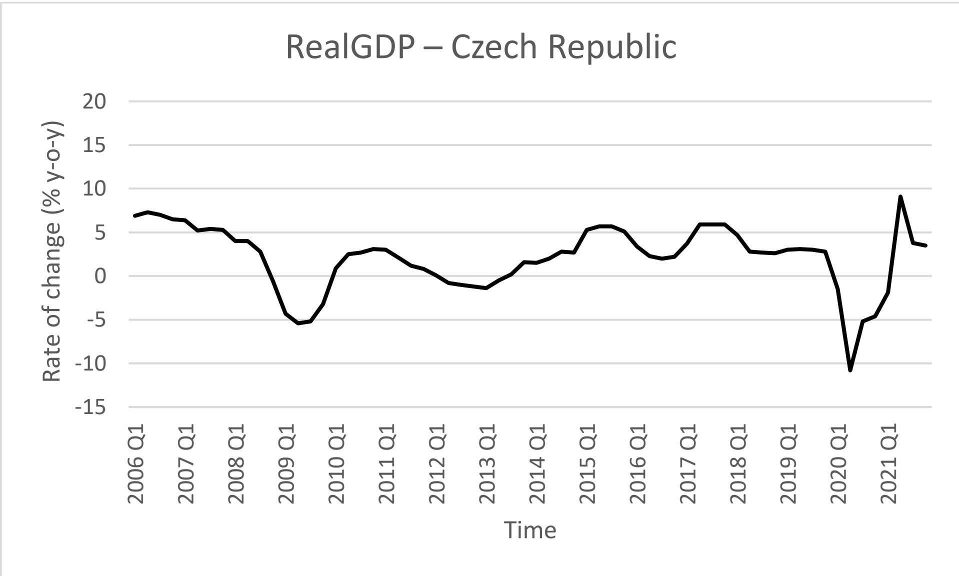


Figure 8.7: Development of variable *M2* in the Czech Republic

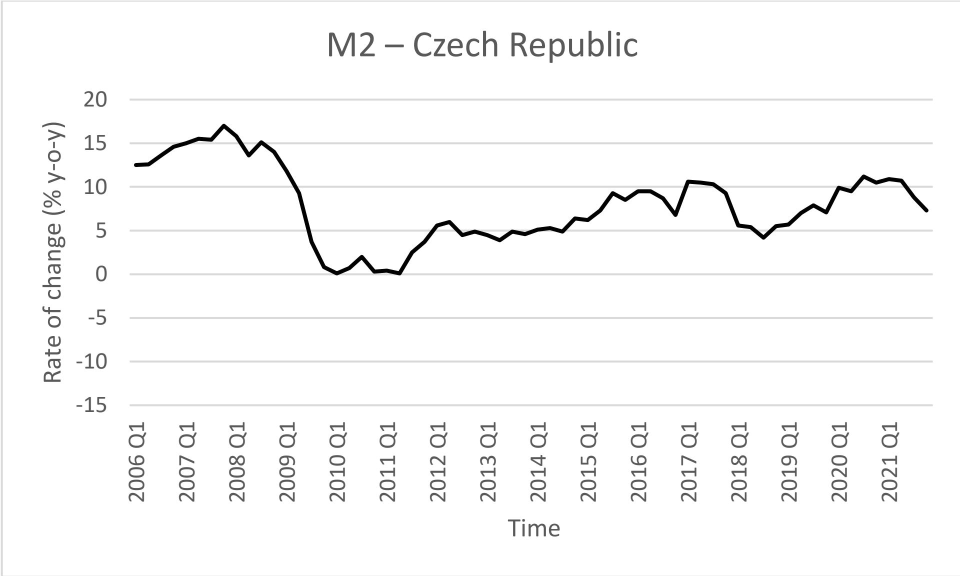


Figure 8.8: Development of variable *NomW* in the Czech Republic

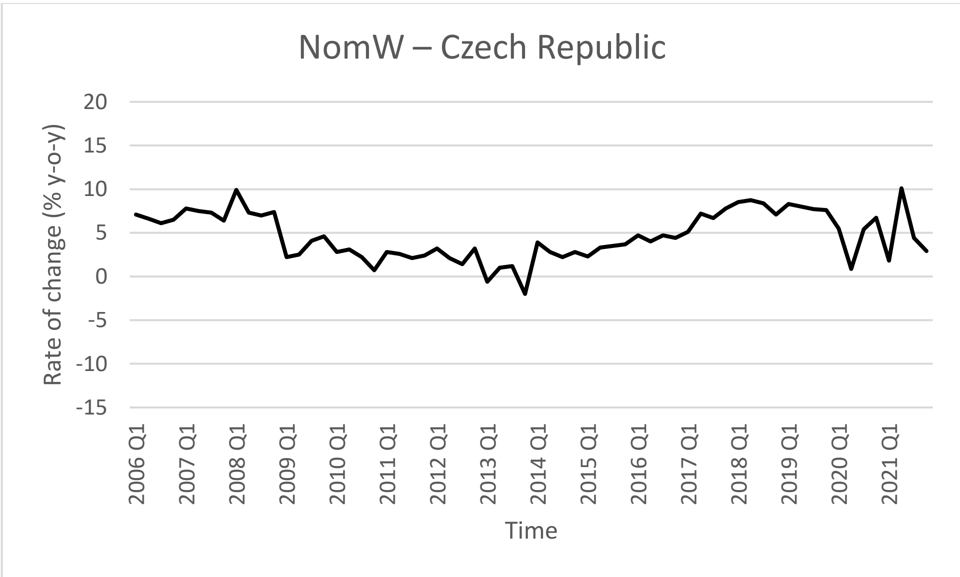


Figure 8.9: Development of variable *NomW* in the Czech Republic in more detail

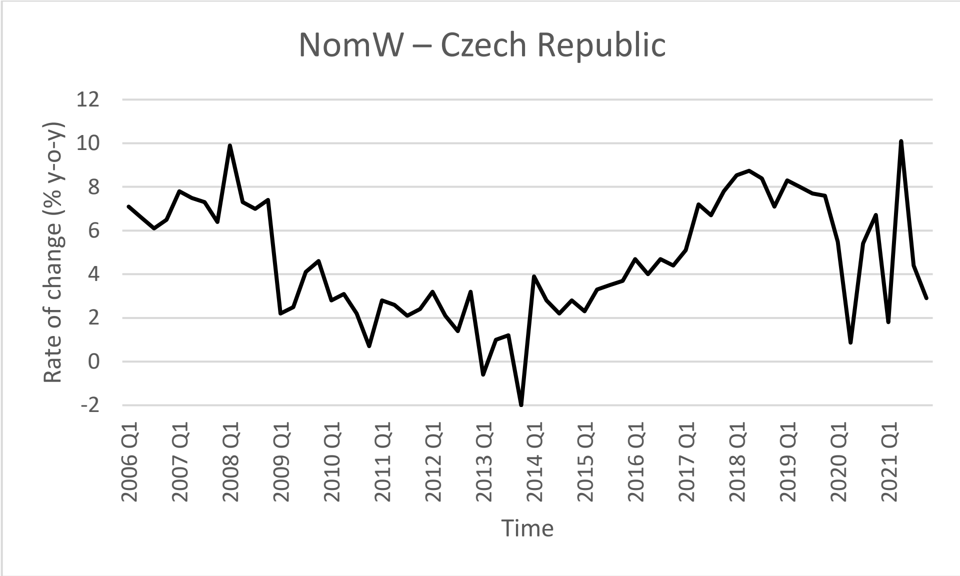


Figure 8.10: Development of variable *RealGDP* in the euro area

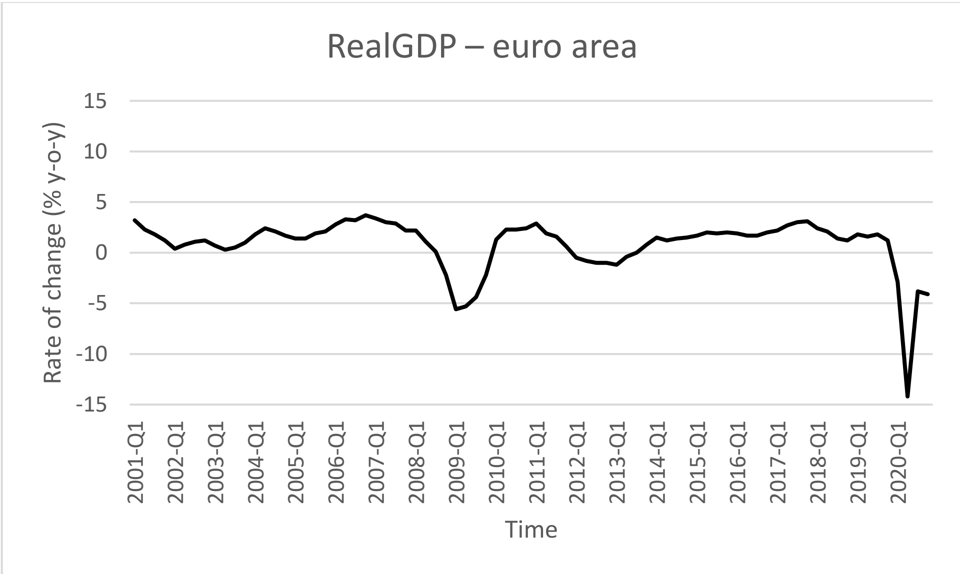


Figure 8.11: Development of variable *M3* in the euro area

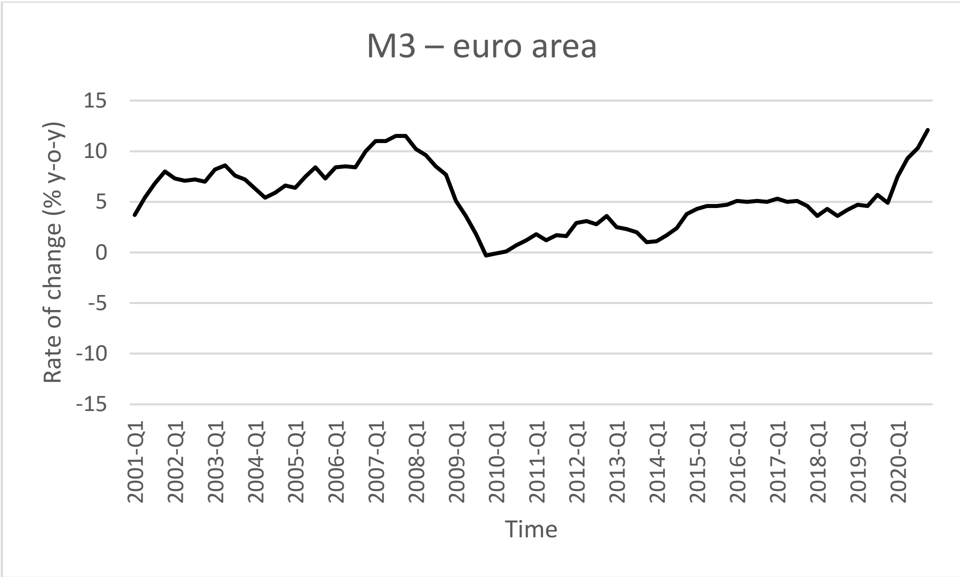


Figure 8.12: Development of variable *LCI* in the euro area

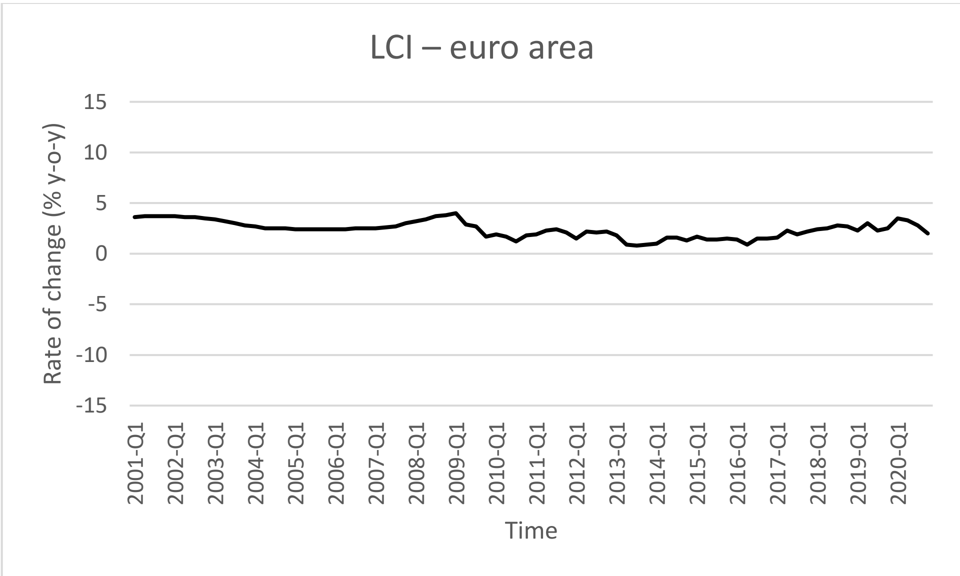


Figure 8.13: Development of variable *LCI* in the euro area in more detail