QUCEH WORKING PAPER SERIES
http://www.quceh.org.uk/working-papers

HAVE WE UNDER-ESTIMATED INFLATION PERSISTENCE BEFORE WW1? US AND INTERNATIONAL EVIDENCE

Stefan Gerlach (EFG Bank)
Rebecca Stuart (Université de Neuchâtel)

Working Paper 25-05

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's Business School
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE

April 2025

Have we under-estimated inflation persistence before WW1? US and international evidence

Stefan Gerlach
EFG Bank and CEPR

Rebecca Stuart
Université de Neuchâtel
and
Queen's University Belfast

April 2025

Abstract

We argue that measurement error in historical price data has led researchers to erroneously believe that there was little persistence of inflation during the 19th century. Using a statistical technique that accounts for these errors, we estimate the persistence of (a) US inflation and (b) inflation in 14 other economies over the period 1842-1913. Our results indicate that persistence approximately doubles when we use this technique.

Keywords: Inflation persistence, gold standard, measurement errors, instrumental variables.

JEL Number: E31, F40, N10.

We are grateful to Peter Kugler, Daniel Kaufmann, Mary O'Sullivan, Morgan Kelly and participants at the World Cliometrics Congress and at the University of Limerick economics seminar for their comments.

Contact information: Rebecca Stuart (corresponding author), email: rebecca.j.stuart@gmail.com, website: https://www.rebeccastuart.net/.

1. Introduction

It is widely believed that consumer price inflation displayed little or no persistence during the 19th century. Studies of the statistical properties of US inflation in the 19th century have generally found that it was close to a white noise process. Indeed, the lack of statistical persistence observed in historical data led Barsky (1987, p. 3) to conclude that, in the US at least, '*Inflation evolved from essentially a white noise process in the pre-World War I years, to a highly persistent, non-stationary ARIMA process in the post-1960 period*'.

One reason for this finding is that the structure of the economy was very different then. Over time, the share of non-durable goods, such as food, in the consumer goods basket has declined. Since their prices tend to display much less persistence than the prices of durable goods and services, one would expect overall inflation rates to be much more sluggish now than a century and half ago.

However, other factors suggest that the conclusion that inflation was not persistent may be mistaken. One such factor is that under the gold standard the money supply was determined by gold in circulation, meaning inflation should have been predictable. Indeed, Barsky and DeLong (1991) show that gold production forecasted inflation during the period of the gold standard. The effect of expected increases in gold in circulation on inflation was widely discussed by the public (Flandreau (2004)). New discoveries of gold attracted particular attention: a column in the Times of London on 25 June 1852[1] concluded: *"we arrive…at the…conclusion that the Californian and Australian discoveries…will produce effects of a momentous character"*, while the effect of gold discoveries on prices was highlighted by the Swiss newspaper Der Bund on 25 January 1858:[2] *"Coins in circulation rose everywhere after the gold findings in California and Australia... The price of all goods rose"*.

A second factor is the well-known problem of measuring historical prices. Indeed, there is a long literature discussing the difficulties constructing historical price indices and measurement errors are thus a likely to have been important.

---

1 https://www.thetimes.co.uk/tto/archive/article/1852-06-25/5/1.html

2 We are grateful to Daniel Kaufmann for this quote which was used in an earlier version of Kaufmann and Stuart (2024). It was obtained via e-newspaperarchives.ch and translated from the original text: "In Folge der Goldausbeute in Kalifornien und Australien [stieg] die Zahl der in Umlauf befindlichen geprägten Münzen [überall], am meisten in Frankreich. […] Der Preis aller Waaren stieg […]"

Take the example of the US. Today, official CPI data are compiled by the Bureau of Labour Statistics. However, no official series was compiled contemporaneously for the period prior to 1913, and so we must rely on indices compiled by various researchers for earlier periods. In a comprehensive study of US cost-of-living data since the 18th century, Officer (2014) identifies a total of 14 different price series covering all or part of the period studied here (1841 to 1913). The existence of so many series is proof of measurement error since at most one can be correct. More likely, we believe, all contain at least some degree of error.

In this paper we investigate the potential role of measurement errors in historical inflation data in biasing down estimates of the degree of inflation persistence. In doing so, we add to the extensive literature by Christina Romer and others which discusses measurement error in historical macroeconomic data such as GNP, industrial production and unemployment.[3] That literature shows, among other things, that failing to account for poor measurement in historical data led researchers to overestimate business cycle volatility in the pre-First World War period. This suggests that, in a similar fashion, measurement error in prices might have led to erroneous conclusions regarding inflation behaviour during the period. Surprisingly, as we discuss further below, few studies have focussed on this issue.

We conduct our analysis using the simplest possible framework, a first-order autoregressive, AR(1), model for inflation. It is well-known that if the regressor in a regression of y on x is subject to white noise measurement errors, the estimated parameter on x will be biased downward. In the case of an AR(1) model, that means that the degree of persistence will be underestimated.

The standard approach to dealing with measurement errors is to use instrumental variables (IV). We do so here. In the first instance, we collect five measures of inflation for the US over the period 1842 to 1913. We use three different instruments to obtain estimates of persistence that are unaffected by measurement error. Our results are remarkably consistent using the three different instruments. On average, the IV estimates of the AR parameter are almost twice as large as simple OLS estimates. Furthermore, Durbin-

3 See for instance, Romer (1986a), (1986b), (1986c), (1989), Miron and Romer (1990), Diebold and Rudebusch (1992).

Hausman-Wu tests indicate that the difference between the OLS and instrumental variables estimates are statistically significant. These findings are what we would expect if measurement errors were present.

We next broaden our analysis using data from 14 countries for which we can compute annual CPI inflation over the period 1842-1913. In this case, we identify an instrument that is unlikely to be correlated with the measurement error in each of the 14 countries. Here, using IV more than doubles the estimated persistence parameters across these countries. Moreover, the IV estimates are statistically different from the OLS estimates.

Overall, we believe that this constitutes strong evidence that measurement errors have caused the persistence in inflation in the 19th century to be underestimated.

We go on to compare our new estimates of persistence in the 19th century with those in more recent data. Although our estimates of historical persistence are significantly larger than previous estimates, the level of persistence remains below that in more recent data. We hypothesise that this may be due to shifts in the consumer basket and provide some simple empirical evidence suggesting that such compositional changes help explain at least a part of the remaining shortfall in inflation persistence.

The rest of the paper is structured as follows. The next section discusses the literature on inflation persistence and measurement errors in price series during this period. Section 3 describes the data. Section 4 discusses the econometric issues arising from measurement error and outlines our strategy for addressing it. Our results for the US are presented in Section 5 and the international evidence is in Section 6. In Section 7 we compare our estimates of historical persistence with estimates for more recent data and Section 8 concludes.

2. Literature review

2.1 Inflation persistence

The properties of inflation in the US during the classical gold standard have been studied by many authors, prompted by the apparent absence of an ex-post Fisher effect in US data in the period before the First World War. Since the Fisher effect states that nominal interest rates should increase one-for-one with *expected* inflation, focus naturally turned to the predictability of inflation. While Fisher (1930) hypothesized that inflation expectations may be a weighted average of current and past inflation, Cagan (1956) argued that the lags required to forecast inflation during that period were simply too long (approximately 10 to 30 years) to be reasonable. Summers (1983) thus concludes that it is difficult to reconcile the data with standard economic models of fully informed and rational agents. Shiller and Siegel (1977) used spectral techniques to identify long- and short-term movements in inflation and interest rates and concluded that inflation was not easily forecasted at the time. Barsky (1987) employed ARIMA models to show that inflation was essentially a white noise process, and therefore unforecastable.

Since these papers were published, inflation persistence has become widely studied. However, there are many definitions of persistence applied in the literature. Fuhrer (2009) distinguishes between "reduced-form" persistence, which is the empirical property of an observed inflation measure, and "structural persistence" which occurs as the result of features of the economy. The empirical regularity of reduced-form persistence was noted first while structural explanations, such as those by Calvo (1983) and Rotemberg (1982, 1983) and subsequent extensions and variations, attempted to map this regularity to the functioning of the economy.

In this paper, we focus on reduced-form persistence. Within the broader inflation literature, reduced-form persistence has been measured in several ways including unit root tests (Barsky (1987), Ball and Cecchetti (1990)), the first-order autocorrelation (Pivetta and Reis (2007), Grytten and Hunnes (2009)), the sum of autoregressive coefficients (Benati (2008)) and unobserved component models that estimate "permanent" and "transitory" components of inflation (Stock and Watson (2007)).

Estimates of the degree of persistence, while all low, vary. Benati (2008) measures persistence as the sum of the autoregressive coefficients of inflation and finds that persistence was entirely absent in inflation data for several countries in the period before 1914. Cogley and Sargent (2015) and Cogley, Sargent and Surico (2015) estimate unobserved component stochastic volatility models of wholesale and consumer price levels in the US and UK, respectively, and find little persistence in inflation in either country during the 19th century. Grytten and Hunnes (2009), using various autoregressive lag lengths across monetary regimes, draw similar conclusions studying data from Denmark, Norway and Sweden during the gold standard.

On the other hand, Meltzer and Robinson (1987) report limited first-order serial correlation in inflation rates during the gold standard in five of the seven economies that they study.[4] Their data are not a balanced panel but cover periods starting between the early-1860s and late-1880s and ending in 1913. Moreover, several papers in the 1990s (Alogoskoufis and Smith (1991), Alogoskoufis (1990), Burdekin and Siklos (1991)) used simple Philips curves or univariate AR processes and often find some persistence during the gold standard.

2.2 Structural change

Changes to the structure of the economy might affect the persistence of inflation. Hanes (1999) demonstrates that the degree of processing of products in a price index can have important implications for its statistical properties. Using producer prices, he shows that the prices of less processed goods are less autocorrelated than those of more processed goods. Since historical data on the cost of living have a greater share of less processed goods than modern data, they will naturally exhibit less persistence.

In a similar manner, the composition of household consumption has changed over time. Oshima (1961) and Juster and Lipsey (1967) argue that a significant shift towards the consumption of durables occurred in the US in the 1920s associated with automobiles and the electrification of household appliances. While this view was challenged by contemporary authors (see Vatter (1967a), (1967b), Vatter and Thompson (1969)), by the time Olney (1990) showed the importance of access to consumer credit in enabling the increase in consumer spending, the issue was largely settled. The increase in demand for durables alongside growth in the service sector, has meant that over the course of the 20th century, the share of non-durables in the consumer basket has declined. Since non-durables includes items such as food and energy for which prices are known to be volatile, this likely affected the inflation persistence of the overall basket.

---

4 The countries are Denmark, Germany, Italy, Japan, Sweden, UK, US. German and Japanese data do not exhibit serial correlation. In addition, serial correlation in Italian data is negative.

2.3 Measurement errors in historical inflation

The fact that new price series, improving on older ones, are regularly produced by economic historians testifies to the fact that we have imperfect measures of historical prices.[5] Kaufmann (2020) and Officer (2014) identify several possible sources of measurement error present in historical US price data. In particular, the scarcity of price quotes for historical periods has several effects by requiring the use of proxies (such as wholesale prices[6] or prices for narrow geographical regions[7]) and interpolations (such as imputing prices of housing and rent[8]). Services prices are scarcer than goods prices, although this in part reflects the fact that services accounted for a smaller share of the economy. The smaller number of price quotes, while reflecting a narrower consumer basket and less variety generally, will tend to raise volatility of historical indices, ceteris paribus. It is notable that these measurement issues relate to the domestic component of inflation, likely making measurement error uncorrelated across countries and implying that international inflation may be a good instrument. We will exploit this feature in our analysis below.

Nonetheless, the effects of measurement error in prices has generally been considered only in passing in the literature. For instance, Benati (2008, p. 1041) states that although *'the problem [of measurement error] is potentially there, unfortunately it is not clear at all how to even gauge an idea of the likely extent of its impact, and in what follows I will therefore ignore it'*. Indeed, we are aware of only a few studies that directly address the issue of measurement errors in inflation.

5 See the case of the US, discussed in Section 4.1, where 14 different price series have been compiled coving subsamples of the period we study.

6 For instance, our UK price data prior to 1870 are from Feinstein (1998), who uses wholesale prices to proxy for the retail prices of flour (1846-1870), pork and bacon (1850-1870), potatoes (1846-1870) and tallow (as a proxy for candles, 1860-1870) (See Appendix to Feinstein (1995) for details). Overall, these four items make up just over a fifth of the total index, during the period that all are used (1860-1870) (See Table 1 in Feinstein (1998)).

7 See, for instance, the discussion below of the series for the US compiled by Adams (1944).

8 For instance, Kaufmann (2020) notes that in his series for the US, Long (1960) approximates the prices of several items, including rent, by a linear interpolation over the entire 1880s, while Lebergott (1964) constructs a reproduction cost index by equally weighting the cost of construction materials and wages for low-skilled workers.

Cogley and Sargent (2015) estimated unobserved component stochastic volatility models of the wholesale price level in the US, allowing for an unobserved measurement error in the price level. They compare the Hanes (2006) modern replication of the Warren and Pearson (1932) historical wholesale price series with the official Bureau of Labor Statistics (BLS) data over the period 1947-1990 and assume that the BLS data are "true" so that the difference between the two series represents "measurement error". Assuming that the statistical properties of this "measurement error" also apply to the 19th century, they find no significant evidence that price level volatility has changed compared to the pre-War period. Cogley, Sargent and Surico (2015) conduct a similar exercise for UK consumer prices while using the same US wholesale price data to inform priors and again find little evidence of a shift in price level volatility in the post-War period. In contrast to these studies, we focus on consumer price inflation in the US and elsewhere, and we do not take a stand on what the "true" price level or inflation rate was during the 19th century.

More recently, Kaufmann (2020) studies the impact of deflation on economic activity. While earlier studies by Atkeson and Kehoe (2004), Bordo and Filardo (2005) and Borio et al. (2015) argue that deflation was only weakly linked to reduced economic activity during the 19th century, Kaufmann (2020) shows that once measurement error is accounted for, the link between deflationary episodes and economic activity is quite strong. He finds large and statistically significant declines in US industrial production growth during periods of deflation. This implies that measurement errors in inflation may have economically meaningful impacts on statistical results.

3. Data

3.1 US data

We start by collecting five measures of US inflation for the sample period 1842-1913 at an annual frequency.[9] They are all intended to capture changes in consumer prices, and the differences between them can be thought of as arising from measurement errors.

---

9 This is the broadest set of price series covering the full sample period 1841 to 1913 that we can find. 1841 is selected as the start date because this is when one of the series (Burgess (1920)) begins.

Officer (2014), which contains an in-depth study of the available historical US cost-of-living data, selected the series by David and Solar (1977)[10] as the best measure of prices for this period. In addition to data quality, Officer bases this judgement on several criteria. First, he favours series that rely to as great an extent possible on retail (rather than wholesale) prices. Second, the coverage of the consumer basket must be as wide as possible and not omit important consumption items such as rent or housing. Third, he prefers series that are calculated using averages of monthly prices over the course of a year, as opposed to, for instance, single observations representing a whole year. Fourth, series using expenditure weights that are as timely as possible are preferred. Fifth, comprehensive geographic coverage is deemed an advantage. Sixth, a sufficiently large number of price series should be used in the calculation of the aggregate series and finally, for the most part, longer series are preferred to reduce the impact of methodological changes. Overall, we follow Officer (2014) in considering David and Solar (1977) the best measure of prices for the period we study.

We think of the series that Officer rejected also as estimates of the true inflation rate, albeit measured with greater error. We collect two other composite series that cover the entire sample period 1842 to 1913 and which are constructed using data series that Officer considered but rejected.[11] These are compiled by Hansen (1925) and the Federal Reserve (1957).[12] The series underlying their construction are listed in Appendix A. Based on the criteria outlined above, Officer judges these series to be substandard for several reasons. In the case of Hansen (1925), which Officer refers to as "out of date", he notes that there is a heavy reliance on wholesale prices in this series, and that rent is missing entirely from the calculation of the series. In the case of the Federal Reserve (1957), the series is made up of a relatively large number of short series, which he considers to be of poor quality.[13]

---

10 As presented in Officer (2014). For the period covered here, this series is the same as one of the two consumer price series ("David-Solar-based") presented in Hanes (2006).

11 The BLS also compiles a composite series for the period, but this overlaps significantly with Brady and Solar. Where it does not (1861-1890), we use the underlying component series (Hoover (1960)) in the calculation of the instruments discussed below (as described in Appendix A).

12 As reported in Bureau of Census (1960).

13 This is true of the entire sample period for which the series is available (1820-1957), however, for the period which we consider, the number of series used is only one greater than the number employed by David and Solar (1977).

In addition, we collect two series that are consistently compiled from a single source for the entire sample period (Burgess (1920) and Adams (1944)).[14] However, these series both have significant drawbacks. The price series of Burgess (1920), which begins in 1842 and thus determines the start-date for our analysis, is based on ten staple food items. The series by Adams (1944), although covering a wider basket of goods including some services, pertains to a very narrow geographic and socioeconomic cohort: prices paid by farmers in Vermont. Nonetheless, they are proxies of prices in the US during this period, albeit measured with error.

3.2 US inflation descriptive statistics

We start by plotting the annual inflation rates of these five series for the period 1842-1913 (Figure 1). While the series move broadly similarly over the sample period, the differences between the series can be quite large: the average range of the series over the entire period is 7 percentage points, and on 14 occasions in the 72-year sample period it is greater than 10 percentage points.

Next, we compute descriptive statistics (Table 1). The average inflation rates range between 0.07% for the Adams series to 0.81% for Burgess. The medians range from -1.08% for the Adams series – indicating many deflationary years in this series – to 0.41% for the Burgess series. The median is zero for the other three series. The standard deviations of the series vary between 4.61 for the Federal Reserve series and 8.22 for the Hansen series.

As this analysis makes clear, there are important differences between the five series. One possible reason is measurement errors that introduce differences in the levels of the series and blow up their variances to varying degrees.

3.3 International inflation data and descriptive statistics

In the second part of our analysis, we use data from 14 economies for which inflation can be computed for the period 1842 to 1913. They are drawn from a variety of sources, indicated in Appendix B, and should be thought of as capturing the cost of living.

---

14 Other series, such as the well-known series compiled by Warren and Pearson (1932), are for producer rather than consumer prices, and are not considered here.

The dashed lines in Figure 2 show the range of inflation across all 14 economies while the solid line shows the median of the series. The difference in inflation is frequently quite large, exceeding 30 percentage points on several occasions, however, this is largely due to outliers. The cross-sectional interquartile range only exceeds 10 percentage points on three occasions, and averages 4.8 percentage points over the sample period.

Descriptive statistics are shown in Table 2. The average annual inflation rate is about 0.38%, and the standard deviation of inflation is around 5.7%. It is notable that no country is a clear outlier. The average pairwise correlation of each series with the other series is also included in Table 2, and indicates a relatively high degree of co-movement, ranging between 0.27 for Finland and 0.53 for Denmark and the UK.

4. Measurement errors

The empirical work in this paper focuses on comparing OLS with IV estimates of the parameter on lagged inflation in a first-order autoregression for inflation. Given the long literature on the problems of measuring prices in the 19th century, we focus on measurement errors as one reason for why they differ.[15]

We consider the simplest possible textbook case of a regression with one mis-measured regressor, for which we have one instrument. We assume that the measurement errors on inflation are normally distributed and serially uncorrelated.[16] We have several data series and index these by i. We assume that the measurement errors are mutually uncorrelated.

We observe the inflation rate as measured by price index i, $\pi_{i,t}$, subject to a measurement error, $\nu_{i,t}$:[17]

(1) $$\pi_{i,t} = \pi^*_{i,t} + \nu_{i,t}$$

15 Another possibility is that in additional to persistent shocks to inflation there are also temporary shocks that are specific to the time series of inflation analysed. In either case, our point remains that the measures of persistence the literature has focussed to date on provide a poor estimate of the underlying inflation dynamics.

16 See for instance the discussion in Greene (2012, Section 8.5).

17 Of course, what we refer to as the measurement error could be some other shock that temporarily raises observed inflation.

where $\pi^*_{i,t}$ denotes the true, unobserved, inflation rate. Consider next a first-order autoregressive model:

(2) $\pi^*_{i,t} = \beta_i \pi^*_{i,t-1} + \varepsilon_{i,t}$

If data on $\pi^*_{i,t}$ were available, we could estimate this equation using OLS. Since they are not, we must rewrite this equation in terms of the observed inflation rate, $\pi_{i,t}$. We have that:

(3) $\pi_{i,t} = \beta_i \pi_{i,t-1} + [\varepsilon_{i,t} - \beta_i \nu_{i,t-1} + \nu_{i,t}]$

where we assume that the regression error $\varepsilon_{i,t}$ is normally distributed and serially uncorrelated. Since the measurement error appears in the composite residual, the error and the regressor are correlated. This is why OLS estimates are biased.

Consistent estimates of β_i can be obtained using instrumental variables. That requires an instrument, z_{t-1}, that is, ideally perfectly, correlated with the unobserved true inflation rate, $\pi^*_{i,t-1}$, but uncorrelated with the measurement error, $\nu_{i,t-1}$. Unfortunately, in applied work it is often difficult to find a good such instrument. We identify three different instruments for the US data and one for the cross-country data. The increase in estimated persistence is remarkably consistent using our different instruments. The use of several instruments in this way strengthens the case for the existence of measurement error. We next discuss these instruments.

4.1 Instruments

First, since we have several measures of US inflation and under our assumption that the measurement errors are uncorrelated across series, the cross-sectional mean and median of the lagged inflation rates may attenuate the measurement error in individual series and are potential instruments. We prefer the median since it is more robust to outliers which, as discussed above, are often present in historical data.

To avoid the problem that the median may lead the instrument to be spuriously significant in the estimated inflation equation proposed above if the cross-sectional dimension is not sufficiently large, we define the "series-specific" median which is the cross-sectional median when all data series are considered, except the series in question.

Thus, the series-specific median for data series i is constructed using all data, except data series i.[18] Below we use the series-specific medians as instruments.[19]

Of course, an instrument must not only be uncorrelated with the measurement error and correlated with the true value of the poorly measured variable, it must also not be an omitted structural variable from the hypothetical regression of the true inflation rate, $\pi_{i,t}^{*}$, on its lag in equation (2). We believe that this is a reasonable assumption in the case of the series-specific median.

The final column of Table 1 shows the correlation between each series and the series-specific median. These correlations are high, ranging from 0.64 for the Hansen series to 0.92 for the David and Solar series.[20]

To see why the median of a set of mis-measured estimates of inflation can be a good instrument, consider for simplicity that the measurement errors have the same variance, σ^2, in all countries and are mutually uncorrelated and consider the cross-sectional mean of inflation. The errors add a term to the variance of the mean equal to σ^2/N , where N is the number of economies. The importance of the measurement errors is thus reduced sharply when a set of series is considered. The same argument applies to the cross-sectional median.[21] However, the assumption that the errors are uncorrelated across inflation estimates for the same country is quite strong, which is why we also use two other instruments.

Second, we use commodity wholesale price inflation as an instrument. Wholesale prices are frequently used as proxies for consumer prices in historical periods as they are considered to be relatively well measured (Kaufmann (2020)). Indeed, wholesale prices are used in some of our inflation series. In some cases, the use of wholesale prices is minimal: the Adams (1944) and Burgess (1920) series rely solely retail prices, while only for the

---

18 The construction of the US instruments is complicated by the composite nature of the series. See Appendix A for details of the underlying series in the US instrument.
19 Pukthuanthong and Roll (2009) use a similar methodology, which they refer to as "out-of-sample principal components", in their study of stock market integration. See also Gerlach and Stuart (2023).
20 In a first order autocorrelation of the difference between each series and the median of the other series, the constant is always insignificant and the AR parameter is insignificant in all cases except for the Hansen series.
21 Wonnacott and Wonnacott (1977, p. 182) state that the variance of the sample median is approximately given by $(\pi/2)(\sigma^2/N)$ and thus declines in the same way as the variance of the mean when N increases.

period 1890 to 1914 are some wholesale prices used alongside retail prices in the series constructed by David and Solar (1977). In the case of the Federal Reserve (1957) series, some wholesale prices are used in the periods 1850 to 1860 and 1890 to 1910. However, the Hansen (1925) series relies on wholesale prices for all non-food items.

Wholesale prices are available for nine commodities for the entire period from the US Bureau of the Census (1975). Since some measurement error in our consumer price series arises from the use of wholesale prices instead of retail prices, to avoid having the same measurement error in our instrument, we select the wholesale prices of commodities which do not proxy for any consumer good.[22] Specifically, we exclude wheat flour, cotton sheeting and coal and use the median of inflation in six commodities: wheat, raw cotton, wool, nails, copper and turpentine. The correlation between each of our five series and our measure of commodity price inflation ranges from 0.47 for the Hansen series to 0.75 for the Adams series.

Third, we can use as an instrument the inflation rate in another country which is likely to co-move with the US inflation rate. While foreign inflation may also be measured with error, that error is unlikely to be correlated with measurement error in the US series. Moreover, it is unlikely that another country's inflation rate would be an omitted variable in a regression of true US inflation on its lag.

Given the gravity theory of international trade, we expect inflation to be transmitted between close neighbours. Indeed, Gerlach and Stuart (2024) show that trade openness and the geographical distance between countries explain much of the variation in the pairwise correlations of inflation between countries during this period. Due to the US's relative remoteness during the period, that makes Canadian inflation a candidate to be an instrument. Moreover, following work by Geloso (2019) and Geloso and Hinton (2020), we believe that Canadian inflation is relatively well-measured.

---

22 It is not possible to verify the exact commodity prices used in the other series. For instance, Hansen (1925b, p.294) notes simply: "For the period 1890 to 1913 an index was constructed consisting of food (40), cloths and clothing (17), fuel and light (6), house furnishing (5)…The food index is the retail prices of the Bureau of Labor Statistics, while the other series are the relative prices at wholesale." However, it seems reasonable to assume that none of the six commodities we include in our instrument would proxy for these categories.

However, Figure 1 shows that there is a substantial increase in the US inflation rate during the Civil War years. This drives a wedge between US and Canadian inflation which peaks at almost 25 percentage points in 1864 and averages over 10 percentage points for the years 1860 to 1865. This is likely to reduce the usefulness of Canadian inflation as an instrument. Indeed, removing the years 1860 to 1865 from the sample increases the average pairwise correlation of the five US series with Canadian inflation from 0.36 to 0.46. We therefore exclude the Civil War years from our sample period in this part of the analysis.

Fourth, for the international analysis, we use the country-specific median as the instrument.[23] That is, when constructing the instrument for country *i* we use data on all countries, except country *i*, thus capturing the common component of inflation across countries. We know that such a component existed in the 19th century since Gerlach and Stuart (2023) estimated reduced-form inflation equations of the form proposed by Ciccarelli and Mojon (2010) and find that the common component is significant in 13 of the 15 economies studied.[24] In addition, we believe that it is unlikely that the lagged country-specific median of inflation in 14 other countries is an omitted variable in a first order autoregression of the true inflation rate in country *i*. Finally, as already noted, we believe that the measurement errors are unlikely to be correlated across countries. Overall, this suggests that it will be a good instrument.

The final column of Table 2 shows the correlation rates between the country inflation rates and country-specific medians. These range from 0.40 for Portugal to 0.80 for Denmark.

23 We do not use the cross-country median as an instrument for the US data as we find that the Civil War and relatively low correlation in inflation between the US inflation rate and that in other countries makes it a weak instrument.

24 The sample period in this study was from 1850 to 1913, and the sample of countries was the same except that Portugal and Spain were not included, but Iceland and Australia were. The latter two are not included here as data prior to 1850 are not available for them.

5. **The effect of measurement error on US inflation persistence**

5.1 Estimates disregarding measurement errors

We first disregard any potential measurement errors and estimate a AR(1) model on data for 1842-1913.

(4) $\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$

The regressions are estimated using OLS. As a robustness check, we present results using seemingly unrelated regressions in Appendix C with overall similar results.

The results are presented in Panel (a) of Table 3. The autoregressive parameter, β_i, is significant at the 5% level in all cases. It is smallest, 0.28, in the case of the Hansen data, which is the series that is most volatile and on average the least correlated with the other series. It is highest, 0.60, in the case of the David and Solar series, which is the second least volatile series and the one we believe is the best measure of inflation. The average parameter estimate across all five regressions is 0.44.

The AR parameter is of crucial interest as it allows us to calculate the persistence of a shock to inflation as measured by the half-life of a shock, which is given by $\frac{ln(0.5)}{ln(\beta_i)}$.[25] This function is nonlinear and steepens as β_i rises. As the AR parameter approaches unity, the slope approaches infinity. For low values of the AR parameter, the half-life is very short. For instance, for a value of 0.1, the half-life is just 0.30 years. In contrast for an intermediate value of 0.5, it is 1; and for a high value of 0.8, it is 6.6 years. This illustrates that to assess correctly the degree of persistence of inflation, it is essential to estimate the AR parameter consistently. Across all five regressions, the average half-life is just 0.88 years.[26] These estimates therefore suggest a very low level of persistence in inflation during the period.

5.2 Estimates using the series-specific median as an instrument

To explore whether the presence of measurement errors may explain the low degree of persistence in inflation, we re-estimate equation (4) using IV. We first use the series-specific median, π_{t-1}^M, as an instrument. In the next section we use inflation in Canada as an instrument.

25 See, for instance, Murray and Papell (2004) and Chortareas and Kapetanios (2004).

26 Since the half-life is a non-linear function of the AR parameter, we obtain the average half-life (in all instances) by averaging the half-life resulting from the different estimates.

Since these estimates are only valid if the instruments are strong, we first regress the inflation rates on their instruments and calculate F-tests of the hypothesis that the slope parameter is zero. Staiger and Stock (1997) propose the rule of thumb that, in the case of a single endogenous variable, if the F-statistic in this first-step regression is greater than 10, the instrument is strong. We report these statistics in Panel (b) of Table 3. The F-statistics range from 48.2 to 391.6. Overall, we conclude that the proposed instruments are "strong" in all cases.[27]

The IV results are presented in Panel (b) of Table 3. The autoregressive parameter is significant in all cases (all p-values are zero). Moreover, the IV estimates of the AR parameter are systematically larger than the OLS estimates. Indeed, the increase in the AR parameter is at times large. For instance, in the case of the Hansen data, the estimated parameter rises from 0.28 to 0.93. At 0.69, the average parameter is over 50% larger than measurement error is ignored. As a result, the average half-life of shocks to inflation rises from 0.88 to 3.26 years, a material difference.

One would expect that the IV estimates of the autoregressive parameter would be larger than the OLS estimates if measurement errors are present. Thus, these results are suggestive of all series being measured with error. Next, we formally test for measurement errors, that is, we test whether the OLS and IV estimates are the same. To do so, we compute Durbin-Hausman-Wu (DHW) tests for simultaneity bias.[28] The test is constructed by performing the first-stage regression implicit in the IV estimation process, that is, by regressing the inflation rates, $\pi_{i,t}$, on the series-specific median of inflation, π_t^M:

(5) $$\pi_{i,t} = \delta_i + \gamma_i \pi_t^M + u_{i,t}$$

The estimated residuals are added to equation (4):

(6) $$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \lambda_i \hat{u}_{i,t-1} + \zeta_{i,t}^*$$

and the hypothesis that $\lambda_i = 0$ is tested using a t-test. The p-values for a test of regressor endogeneity are also included in Table 3. They reject the hypothesis that the parameters in the OLS and IV regressions are the same in three cases (Burgess, the Federal Reserve and

27 More recently, Lee et al., (2022) propose alternative critical values for single-IV models that depend on the value of the first-stage F-statistic. The results remain significant at the 5% level using these critical values.

28 See Greene (2007, p 276).

Hansen). Since the David and Solar series is arguably the series measured with the smallest error, it is perhaps less surprising that the null hypothesis is not rejected in this case, although a similar argument cannot be made for the Adams series.

To increase power, we perform these tests jointly, rejecting the null hypothesis that the two estimates are the same (*p*-value 0.00).

5.3 Estimates using commodity price inflation as an instrument

The results using commodity price inflation as an instrument are included in Panel (c) of Table 3. The first stage F-statistics range from 19.97 in the case of the Hansen series to 81.41 for the Adams series.[29] The IV estimates indicate that the AR parameter is significant at the 1% level in all cases and ranges from 0.81 (Adams series) to 1.16 (Hansen series), and averages 0.97. This is markedly higher than the average OLS estimate (0.44), and indeed DHW tests indicate that the IV estimates are significantly different from the OLS estimates (in all cases, the *p*-value = 0.00). A joint DHW test also returns a *p*-value of 0.00.

5.4 Estimates using Canadian inflation as an instrument

We next turn to the results using our third instrument, Canadian inflation. However, before proceeding, we will jump ahead somewhat to note that while using the Hansen data, the first-stage F-statistic in the IV analysis is 5.99, indicating a weak instrument. We therefore only report results for the four other series for which the first-stage F-statistic is greater than 10.

The OLS results excluding the Civil War period are presented in Table 4. The AR parameter estimates are statistically significant at the 5% level in the case of the Federal Reserve and Brady and Solar series, and at the 10% level in the case of the other two series. However, the average of the parameter estimate is smaller than the full sample estimates that incorporate the observations for the Civil War (0.44 compared to 0.22).

Turning to the IV results (middle Panel of Table 4), the AR parameter estimates are all statistically significant at the 5% level.[30] They are also substantially higher than the OLS estimates, at least doubling in size in all cases. Indeed, the average parameter estimate

29 The results in this section are unaffected by applying the Lee et al., (2022) critical values.
30 Using the Lee et al., (2022) critical values, we find that the result for the Federal Reserve series is not significant at the 5% level.

across all four series is 0.92. Finally, p -values from DHW tests (bottom of Table 4) indicate that the IV parameter estimates are significantly different from the OLS estimates at the 5% level in three of the four cases, the exception being the Federal Reserve series for which the p -value is 0.06. Unsurprisingly, a joint DHW test returns a p -value of 0.00.

While one can always make an argument against a specific instrument, our results are remarkably consistent across the three instruments we use here. Taking the analysis from all three instruments together, on average, the estimated AR parameter when IV is used is 0.85, close to double the OLS estimate of 0.44 (0.24 when the Civil War is dropped).[31] Overall, we believe that the consistency of the results across our estimates provides strong evidence that measurement error has biased downward estimated persistence in US inflation in the period before the First World War.

6. International evidence on inflation persistence

The results above suggest that measurement error in inflation rates may be the reason why inflation persistence is estimated to be so limited in the US before 1913. We next explore whether this is also the case in other economies, using data from 14 countries for which we can compute annual inflation over the period 1842-1913.

As before, we first discuss the OLS estimates before turning to the IV estimates, the critical questions of the power of our instrument and whether there is evidence of simultaneity bias.

6.1 Estimates disregarding measurement

Our approach is the same as in Section 5. We first estimate equation (4) where the subscript i now refers to the inflation rate in economy i. The results are available in Table 5. The AR parameter, β_i, is significant in the cases of Canada, Denmark, Germany, Netherlands, Norway, Sweden and the UK, that is, in 7 of 14 regressions. It ranges between -0.12 in Spain to 0.37 in Denmark. On average it is 0.18. Since the half-life of shocks can only be computed for AR parameters that are between zero and unity, we disregard negative parameters (Portugal and Spain) and compute the average half-life, which is 0.47 years. As in the case

---

31 Ignoring the estimates that exclude the Civil War period, the average estimated parameter is 0.83.

of the US OLS results in Section 5, these results also suggest that there is little persistence to inflation.

6.2 Estimates using the country-specific median as an instrument

Next, we re-estimate equation (4) using IV and the country-specific mean as an instrument. We report the first stage F-statistics in the bottom part of Table 5. They range from 14.6 in Portugal to 152.4 in the UK. The results of our IV estimation are in the middle Panel of Table 5. The autoregressive parameter is significant in 9 of 14 cases – Belgium, Denmark, Finland, Germany, the Netherlands, Norway, Sweden, Spain and the UK.[32] Thus, we cannot reject the hypothesis of no persistence in several cases. However, since we have several equations, it seems sensible to test the joint restriction that all the β_i parameters are zero. This yields a p-value of 0.000.[33]

Moreover, the parameter estimates, which range from -0.10 in Switzerland to 1.04 in Finland, average 0.43 across all regressions, that is, more than twice the average OLS estimate. As a result, the average half-life is 0.93 of a year, that is, about twice the average when OLS is used.

6.3 Are the IV estimates significantly larger?

While the point estimates of the AR parameter frequently double or triple when IV is used compared to OLS, are the two sets of parameters significantly different from each other? Note that if there are no measurement errors, then the OLS and IV estimates are both consistent and we would expect them to differ randomly. Yet, in 12 of 14 cases the IV estimates are larger than the OLS estimates, the probability of which is less than 1% if they had the same expected value. This suggests that measurement errors are present.

To assess whether the OLS and IV estimates of β_i are statistically different from each other, we compute Durbin-Hausman-Wu (DHW) tests for simultaneity bias. The p-values (bottom of Table 5) show that the hypothesis that the parameter is zero can be rejected in 6 of 14 cases at 5% level (and can be rejected in a further 2 at the 10% level). That casts some doubt on the measurement error hypothesis.

32 Using the Lee et al., (2022) critical values, we find that the results for Finland and Sweden are not significant at the 5% level.

33 Similarly, a joint test of the OLS estimates leads to a p-value of 0.000.

However, that finding may be due to a lack of power of the IV estimates. We therefore perform a test of the joint DHW test. That test yields a p -value of 0.00. Overall, we conclude that, when considering the results for the 14 economies together, the parameter is larger when IV is used.

Finally, since instrumental variable estimates are valid asymptotically, as a robustness test, in Appendix D we re-estimate our results using data for the period from 1800 to 1913 for a sub-sample of countries and US series for which we have data for this extended period. The results support our findings here. Taken together, these results are compatible with inflation being measured subject to errors.

6.4 Comparing estimated persistence in US and international data

Table 6 compares the estimated AR parameters using OLS and IV for US and international inflation data. Ignoring measurement error, the average OLS estimate of the AR parameter in the US data is 0.44 (0.24 when the Civil War is dropped), and in the non-US data, it is 0.18.

When using IV, the average AR parameter estimate for the US series is 0.69 when the series-specific median is used as an instrument, 0.97 when the median commodity price inflation is used as instrument and 0.92 when Canadian inflation is used (and the Civil War period is dropped). The latter two averages are influenced by the estimated AR parameter in excess of 1 in the regressions for the Hansen and Federal Reserve series, respectively. Excluding these estimates, the average parameter estimate are 0.93 and 0.77, respectively. In the non-US data, the average parameter estimates when IV is used is lower, 0.44.

Thus, when IV is used, the estimated AR parameter increases between 150% and 320% compared to OLS, and over 200% on average. Overall, while the choice of instruments can always be challenged, we view the consistency of our findings across all four instruments as compelling evidence that measurement error has led researchers to underestimate the level of persistence in inflation in the 19th century.

7. Was historical inflation as persistent as it is today?

Does this mean that inflation was as persistent in the 19th century as it is today? Figure 3 presents the estimated parameters from a simple first order autoregression of inflation using data from 1960 to 2020 for the US and the 14 countries that we study above.[34] On average, the OLS estimate of the AR parameter is 0.84. That is somewhat higher than most of our IV estimates on the historical data.

What factors may explain these differences? We consider one important structural reason why inflation persistence might be lower in the 19th century than today: the shifting roles of durables, non-durables and services in households' spending patterns.

Figure 4 shows the estimated share of non-durables in household consumption in the US from 1839 to 2020. The overall trend is clear. The share declined somewhat from 69% of overall spending in 1839 to 61% by 1914. However, in the 20th century, and particularly following the durables "revolution" in the 1920s, the decline in the share of non-durables is marked: by 2020, non-durables accounted for just 22% of overall consumer spending.

Since it is well known that non-durables prices, which include items such as food and energy, are more volatile than durables and services prices, this is one structural reason why inflation might be more persistent today than in the past. Indeed, estimating the persistence parameter by sector on US data for the period 1960 to 2020, we find that it is about 0.85 in the cases of durables and services, which have grown in importance over time. For non-durables, however, it is merely 0.61. Of course, focussing solely on non-durable goods is misleading since historical price indices include some durables and services. However, quantitatively, these estimates of persistence are broadly similar to those in the historical period. This suggests to us that some combination of measurement error and structural factors may explain the lower level of persistence during the Gold Standard.

8. Conclusion

In this paper we have argued that the contradictory findings in the literature that estimates the degree of inflation persistence in 19th century data with literature in economic history that argues that the operation of the gold standard and new gold discoveries led to predictable movements in inflation may be due to inflation being measured with error.

34 Sample period 1955-2020. For sources, see Appendix B.

Much of the existing literature fails to engage with the issues of measurement error when estimating persistence in historical inflation data. We use instrumental variables to deal with the well-known problem of attenuation in parameter estimates in the presence of measurement error. We carry out our analysis on two sets of data. First, we explore the issue of measurement error in US inflation before generalising our results by studying inflation in 14 economies. While the merits of an individual instrument can always be questioned, our results are remarkably consistent, suggesting that using instrumental variables returns parameter estimates that are on average almost double or triple those obtained using simple OLS. Moreover, Durbin-Hausman-Wu tests indicate that the difference between the OLS and instrumental variables estimates are statistically significant.

Overall, we have shown that measurement error is an important likely reason for the lack of statistical evidence of persistence in historical inflation.

References

Abildgren, K., (2009) "Consumer Prices in Denmark, 1502-2007," Working paper 60, Danmarks Nationalbank.

Adams, T.M., (1944), Prices Paid by Vermont Farmers for Goods and Services and Received by Them for Farm Products, 1790-1940; Wages of Vermont Farm Labor, 1780-1940. Burlington, University of Vermont and State Agricultural College, Vermont Agricultural Experiment Station.

Alogoskoufis, G.S., (1992), "Monetary Accommodation, Exchange Rate Regimes and Inflation Persistence", *Economic Journal*, 102(412), 461-480.

Alogoskoufis, G.S. and R. Smith, (1991), "The Phillips Curve, The Persistence of Inflation, and the Lucas Critique: Evidence from Exchange-Rate Regimes", *American Economic Review*, 81(5), 1254- 1275.

Atkeson, A. and P.J., Kehoe, (2004), "Deflation and depression: Is there an empirical link?", *American Economic Review*, 94(2), 99–103.

Ball, L. and S.G. Cecchetti, (1990), "Inflation and uncertainty at short and long horizons", Brookings Papers on Economic Activity, 215-254.

Barquín, R., (2001), Precios de trigo e índices de consumo en España, 1765-1883, Universidad de Burgos.

Barsky, R.B. (1987), "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," Journal of Monetary Economics, 19, 3-24.

Barsky, R.B. and B.J. DeLong, (1991), "Forecasting pre-World War 1 inflation: the Fisher effect and the Gold Standard", *Quarterly Journal of Economics*, August, 106(3), 815-836.

Benati, L., (2008) "Investigating Inflation Persistence Across Monetary Regimes," The Quarterly Journal of Economics, 123(3), 1005-1060.

Bertram, G.W. and M.B. Percy (1979), "Real Wage Trends in Canada, 1900-26: Some Provisional Estimates", *Canadian Journal of Economics*, 12, 299-312.

Bordo, M.D. and A. Filardo (2005), "Deflation in a historical perspective," BIS Working Papers No 186, November.

Bordo, M.D. and A.J. Schwartz, (1988), "Transmission of Real and Monetary Disturbances under Fixed and Floating Rates", *Cato Journal*, 8(2), 451-475.

Borio, C., M. Erdem, A. Filardo, and B. Hofmann, (2015), "The costs of deflations: A historical perspective", *BIS Quarterly Review*, March, 31-54.

Burdekin, R.C.K. and P.L. Siklos, (1999), "Exchange Rate Regimes and Shifts in Inflation Persistence: Does Nothing Else Matter?", *Journal of Money, Credit and Banking*, 31(2), 235-247.

Bureau of Census, (1960), Historical Statistics of the United States, Colonial Times to 1957, U.S. Government Printing Office, Washington DC.

Burgess, W.R., (1920), Trends of School Costs, Russell Sage Foundation, New York.

Cagan, P., (1965), Determinants and Effects of Changes in the Supply of Money, National Bureau of Economic Research: New York.

Calvo, G., (1983), "Staggered prices in a utility-maximizing framework", *Journal of Monetary Economics*, 12(3), 383-398.

Chortareas G. and G. Kapetanios, (2004), "How Puzzling is the PPP Puzzle? An Alternative Half-Life Measure of Convergence to PPP", Queen Mary University of London Working Paper No. 522.

Ciccarelli, M. and B. Mojon (2010), "Global Inflation," *Review of Economics and Statistics*, 92, 524-535.

Cogley, T. and T.J. Sargent (2015), "Measuring price-level uncertainty and instability in the United States, 1850-2012", *The Review of Economics and Statistics*, 97(4), 827-838.

Cogley, T., T.J. Sargent and P. Surico, (2015), "Price-level uncertainty and instability in the United Kingdom", *Journal of Economic Dynamics and Control*, 52, 1-16.

Commission on the Necessaries of Life, (1920), Report, Wright and Potter, State Printers, Boston.

David, P.A., and P. Solar, (1977), "A Bicentenary Contribution to the History of the Cost of Living in America.", in Paul Uselding, ed., Research in Economic History, vol. 2, 1-80.

Douglas, P.H., (1926), "The Movement of Real Wages and Its Economic Significance", *American Economic Review*, 16, 17-53.

Diebold, F.X. and G.D. Rudebusch, (1992), "Have Postwar Economic Fluctuations Been Stabilized?", *American Economic Review*, 82(4), 993-1005.

Edvinsson, R. and J. Söderberg (2010), "The evolution of Swedish consumer prices 1290–2008," in Historical Monetary and Financial Statistics for Sweden: Exchange rates, prices, and wages, 1277–2008, edited by Edvinsson, Tor Jacobson, and Daniel Waldenström, Sveriges Riksbank.

Federal Reserve Bank of New York, (1957), "Index of Estimated Cost of Living in the United States", Research Department, mimeo.

Federal Reserve Economic Data (FRED), "Consumer Price Index in the United Kingdom," https://fred.stlouisfed.org/series/CPIUKA

Feinstein, C.H., (1998), "Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution", *Journal of Economic History*, 58(3), 625-658.

Feinstein, C.H., (1995), “Changes in nominal wages, the cost of living and real wages in the United Kingdom over two centuries, 1780-1990” in Scholliers, P. and V. Zamagni (eds), Labour’s reward: real wages and economic change in 19th and 20th century Europe, Edward Elgar, England.

Fisher, I., (1930), The Theory of Interest, New York: MacMillan.

Flandreau, M., (2004), The Glitter of Gold: France Bimetallism, and the Emergence of the International Gold Standard, 1848-1873, Oxford University Press, England.

Fuhrer, J.C., (2010), “Inflation Persistence”, in Friedman, B.M and M. Woodford, eds., Handbook of Monetary Economics, vol. 3, 423-486.

Gallman, R.E., (1966), “Gross National Product in the United States, 1834–1909”, in D.S: Brady, ed., Output, Employment, and Productivity in the United States after 1800, NBER.

Geloso, Vincent (2019), “A price index for Canada, 1688 to 1850”, *Canadian Journal of Economics*, 52, 526–560.

Geloso, V. and M. Hinton (2020), “Improving deflators or estimating Canadian economic growth, 1870–1900”, *Research in Economic History*, 36, 125–150.

Gerlach, S. and R. Stuart, (2023), “International Co-movements of Inflation, 1851-1913”, *Oxford Economic Papers*, 1-17.

Gerlach, S. and R. Stuart, (2024), “The Geography of Inflation, 1851-1913”, *mimeo*.

Greene, W.H., (2007), Econometric Analysis,6th ed., Pearson Education Limited, England.

Greene, W.H., (2012), Econometric Analysis, 7th ed., Pearson Education Limited, England.

Grytten, O.H. (2004), “A Consumer Price Index for Norway, 1516-2003,” in Historical Monetary Statistics for Norway 1819–2003,” in Eitrheim, Øyvind, Jan T. Klovland and Jan F. Qvigstad (eds.), Historical Monetary Statistics for Norway 1819–2003, Occasional Papers 35, Norges Bank. Downloaded from https://static.norges-bank.no/contentassets/de6b294d63d94b1a859408c33c1b3b5a/hele.pdf?v=03/09/2017122252&ft=.pdf

Grytten, O.H. and A. Hunnes, (2009), “Price Stability and Inflation Persistence During the International Gold Standard: The Scandinavian Case”, Norges Bank Working Paper No. 2009/20.

Hanes, C., (1999), “Degrees of Processing and Changes in the Cyclical Behavior of Prices in the United States, 1869-1990”, *Journal of Money, Credit and Banking*, 31(1), 35-53.

Hanes, C., “Prices,” in S. B. Carter, S. G. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, eds., Historical Statistics of the United States: Earliest Times to the Present, Millennial Edition, vol.3, Cambridge University Press.

Hansen, A.H., (1925), "Factors Affecting the Trend of Real Wages", *American Economic Review,* 15(1), 27-42.

Heikkinen, S., (1997), "Labour and the Market. Workers, Wages and Living Standards in Finland, 1850-1913". Commentationes Scientiarium Socialium 51, Helsinki.

Hoover, E.D., (1960), "Retail Prices after 1850", in Studies in Income and Wealth ed., Trends in the American Economy in the Nineteenth Century, vol. 24, 141-190, NBER, Princeton University Press.

Historical Statistics of Switzerland, 2012. Table H.17. https://hsso.ch/2012/h/17

Instituto Nacional de Estatística, (2001), Portuguese Historical Statistics.

Juster, F.T. and R.E. Lipsey, (1967), "A Note on Consumer Asset Formation in the United States", *Economic Journal*, 77(308), 834-847.

Kackmeister, A., (2007), "Yesterday's Bad Times Are Today's Good Old Times: Retail Price Changes Are More Frequent Today than in the 1890s", *Journal of Money, Credit and Banking*, 39(8), 1987-2020.

Kapteyn, A., and D.G. Fiebig, (1981), "When are two-stage and three-stage least squares estimators identical?", *Economics Letters*, 8(1), 53-57.

Kaufmann, D., (2020), "Is deflation costly after all? The perils of erroneous historical classifications," *Journal of Applied Econometrics*, 35, 614-628.

Kaufmann, D. and R. Stuart, (2024), "Private Money and Money Market Integration: The Role of Payments Infrastructure in 19th Century Switzerland," IRENE Working Papers, 24-05.

Kuznets, S., (1961), "Capital in the American Economy: Its Formation and Financing", in S. Kuznets and E. Jenks, eds., Capital in the American Economy: Its Formation and Financing, NBER.

Lee, D.S., J. McCrary, M.J. Moreira and J. Porter, (2022), "Valid t -Ratio Inference for IV", *American Economic Review*, 112(10), 3260-3290.

Lebergott, S., (1964), Manpower in economic growth: The american record since 1800, New York: McGraw-Hill.

Levin, A. and J. Piger, (2004), "Is inflation persistence intrinsic in industrial economies?", European Central Bank working paper No. 334.

Long, C.D., (1960), Wages and Earnings in the United States, 1860-1890, Princeton: Princeton University Press.

Meltzer, A and S. Robinson, (1989), "Stability Under the Gold Standard in Practice", in M.D. Bordo, ed., Money, History, and International Finance: Essays in Honor of Anna J. Schwartz, NBER, University of Chicago Press.

Miron, J.A. and C.D. Romer, (1990), "A New Monthly Index of Industrial Production, 1884-1940", *Journal of Economic History*, 50(2), 321-337.

Mitchell, B. R. (2003), International Historical Statistics: Europe 1750 – 2000, Palgrave Macmillan, 5TH ed.

Mitchell, W.C., (1908), Gold, Prices, and Wages under the Greenback Standard, Berkeley University Press.

Mühlpeck, V., Sandgruber, R. and Woitek, H. (1979) Index der Verbraucherpreise 1800-1914, in Geschichte und Ergebnisse der zentralen amtlichen Statistik in Österreich 1829-1979, *Österreichisches Statistisches Zentralamt*, 123–67.

Murray, C.J. and D.H. Papell, (2004), "The Purchasing Power Parity Puzzle is Worse Than You Think", University of Houston, *mimeo*.

Officer, L.H., (2014), 'What Was the Consumer Price Index Then? A Data Study, retrieved from www.measuringworth.com/docs/cpistudyrev.pdf

Olney, M.L., (1990), "Demand for consumer durable goods in 20th century America", *Explorations in Economic History*, 27(3), 322-349.

Oshima, H.T., (1961), "Consumer Asset Formation and the Future of Capitalism", *Economic Journal*, 71(281), 20-35.

Pivetta, F. and R. Reis (2007), "The persistence of inflation in the United States", *Journal of Economic Dynamics and Control*, 31, 1326-1358.

Pukthuanthong, K. and R. Roll (2009), "Global market integration: An alternative measure and its application", *Journal of Financial Economics*, 94, 214-232.

Ravenna, F. and M.M. Ingholt, (2021), "The impact of inflation targeting: Testing the good luck hypothesis", *Canadian Journal of Economics*, 54(1), 443-470.

Rees, A., (1961), Real Wages in Manufacturing, 1890-1914, NBER, Princeton University Press.

Romer, C.D., (1986a) "Spurious Volatility in Historical Unemployment Data," *Journal of Political Economy*, 94(1), 1-37.

Romer, C.D., (1986b) "Is the Stabilization of the Postwar Economy a Figment of the Data?" *American Economic Review*, 76(3), 314-34.

Romer, C.D., (1986c) "New Estimates of Prewar Gross National Product and Unemployment," *Journal of Economic History*, 46(2), 341-52.

Romer, C.D., (1989), "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908", *Journal of Political Economy*, 97(1), 1-37.

Rotemberg, J.J., (1982), "Sticky prices in the United States", *Journal of Political Economy*, 90(6), 1187-1211.

Rotemberg, J.J., (1983), "Aggregate consequences of fixed costs of price adjustment", *American Economic Review*, 73(3), 433-436.

Shiller, R.J. and J.J. Siegel, (1977), "The Gibson Paradox and Historical Movements in Real Interest Rates", *Journal of Political Economy*, 85(5), 891-908.

Staiger, D. and J.H. Stock, (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3), 557-586.

Stock, J.H. and M. Watson, (2007), "Has inflation become harder to forecast?", *Journal of Money, Credit, and Banking*, 39, 3-34.

Summers, L., (1983), "The nonadjustment of nominal interest rates: A study of the Fisher effect.", in James Tobin, ed., Macroeconomic prices and quantities: Essays in memory of Arthur Okun, Brookings Institution, Washington, DC.

U.S. Senate, (1893), "Wholesale Prices, Wages, and Transportation, Part 2", Report 1394, 52nd Congress, 2nd session, Government Printing Office, Washington DC.

Vatter, H.G., (1967a), "Has There Been a Twentieth-Century Consumer Durables Revolution?", *Journal of Economic History*, 27(1), 1-16.

Vatter, H.G., (1967b), "Consumer-Asset Formation in the United States: A Rejoinder", *Economic Journal*, 77(308), 848-855.

Vatter, H.G. and R.L. Thompson, (1966), "Consumer Asset Formation and Economic Growth--The United States Case", *Economic Journal*, 76(302), 312-327.

Warren, G.F. and F.A. Pearson, (1932), "Wholesale Prices in the United States for 135 Years, 1797 to 1932", Cornell University Agricultural Experiment Station, Memoir 142.

Wonnacott, T.H., and R.J. Wonnacott (1977), Introductory Statistics for Business and Economics, Second ed., John Wiley & Sons, New York.

Table 1: Descriptive statistics, inflation series for the US, 1842-1913

Series	Average annual inflation rate	Median annual inflation rate	Standard deviation	Average pairwise correlation with other four series	Correlation with median of the other series
Adams (1944)	0.07	-1.08	6.45	0.74	0.89
Burgess (1920)	0.81	0.41	5.84	0.69	0.85
Federal Reserve (1957)	0.71	0.00	4.84	0.66	0.69
Hansen (1925)	0.29	0.00	8.00	0.63	0.64
David and Solar (1977)	0.18	0.00	5.23	0.77	0.92

Table 2:

Descriptive statistics, inflation rates in 14 countries, 1842-1913

Country	Average annual inflation rate	Median annual inflation rate	Standard deviation	Average pairwise correlation with other 13 series	Correlation with median of the other series
Austria	0.90	0.76	3.76	0.30	0.44
Belgium	0.04	0.00	5.36	0.42	0.59
Canada	-0.18	0.15	7.11	0.34	0.48
Denmark	0.28	0.70	3.87	0.55	0.80
Finland	0.37	0.05	6.21	0.28	0.41
France	0.35	0.00	1.77	0.38	0.56
Germany	1.31	1.30	8.93	0.50	0.75
Netherlands	-0.10	-0.33	4.87	0.54	0.78
Norway	0.59	0.68	4.31	0.52	0.75
Portugal	0.70	0.00	9.20	0.32	0.40
Spain	0.17	-0.54	6.72	0.31	0.47
Sweden	0.61	0.49	5.13	0.45	0.61
Switzerland	0.47	0.90	9.01	0.47	0.70
UK	-0.02	0.00	4.38	0.55	0.79

Sources: See Appendix B.

Table 3: OLS and IV estimates of inflation equation, 1843-1913

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Five US inflation series, two instruments

Dependent Variable	**Adams (1944)**	**Burgess (1920)**	**Federal Reserve (1957)**	**Hansen (1925)**	**David and Solar (1977)**
		(a) OLS results			
Constant	0.12 (0.69) [0.17]	0.46 (0.64) [0.72]	0.49 (0.50) [0.97]	0.36 (0.92) [0.39]	0.19 (0.49) [0.38]
Lagged inflation	0.44 (0.11) [4.08]**	0.41 (0.11) [3.71]**	0.47 (0.10) [4.62]**	0.28 (0.11) [2.43]*	0.60 (0.09) [6.35]**
Observations:	71	71	71	71	71
R-squared:	0.19	0.17	0.24	0.08	0.37
		(b) IV results using the series-specific median as an instrument			
Constant	0.11 (0.70) [0.16]	0.27 (0.67) [0.40]	0.26 (0.54) [0.49]	0.20 (1.11) [0.18]	0.19 (0.49) [0.38]
Lagged inflation	0.50 (0.12) [4.11]**	0.68 (0.14) [5.03]**	0.77 (0.16) [4.92]**	0.93 (0.22) [4.27]**	0.57 (0.10) [5.59]**
Observations:	71	71	71	71	71
First-stage F-statistic	256.25	217.53	59.35	45.86	379.40
DHW test *p*-value	0.28	0.00	0.00	0.00	0.52
		(c) IV results using commodity price inflation as an instrument			
Constant	0.09 (0.75) [0.12]	0.04 (0.78) [0.05]	0.11 (0.60) [0.19]	0.15 (1.25) [0.12]	0.14 (0.53) [0.26]
Lagged inflation	0.81 (0.16) [5.16]**	1.00 (0.23) [4.39]**	0.98 (0.22) [4.42]**	1.16 (0.32) [3.61]**	0.91 (0.15) [5.93]**
Observations:	71	71	71	71	71
First-stage F-statistic	81.41	33.81	28.02	19.97	51.12
DHW test *p*-value	0.00	0.00	0.00	0.00	0.00

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.

Table 4: OLS and IV estimates of inflation equation, 1843-1860, 1966-1913

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Four US inflation series, Canadian inflation as the instrument

Dependent Variable	**Adams (1944)**	**Burgess (1920)**	**Federal Reserve (1957)**	**David and Solar (1977)**
		OLS results		
Constant	-0.62	-0.01	0.07	-0.52
	(0.53)	(0.54)	(0.45)	(0.39)
	[-1.16]	[-0.02]	[0.16]	[-1.33]
Lagged inflation	0.22	0.22	0.30	0.26
	(0.12)	(0.12)	(0.11)	(0.12)
	[1.86]	[1.82]	[2.67]**	[2.21]*
Observations:	66	66	66	66
R-squared:	0.05	0.05	0.10	0.07
		Instrumental variables results		
Constant	-0.19	-0.03	-0.01	-0.04
	(0.61)	(0.78)	(0.51)	(0.49)
	[-0.31]	[-0.03]	[-0.03]	[-0.07]
Lagged inflation	0.69	1.25	0.77	0.85
	(0.22)	(0.44)	(0.32)	(0.25)
	[3.11]**	[2.85]**	[2.42]*	[3.35]**
Observations:	66	66	66	66
		Instrumental variables diagnostic tests		
First-stage F-statistic	32.23	17.22	12.26	27.19
DHW test *p*-value	0.00	0.00	0.06	0.00

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively. The Hansen (1925) series is not included in the analysis since the first-stage F-statistics is less than 10.

Table 5
OLS and IV estimates, 14 economies, 1843-1913, of inflation equation

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Dependent variable:	Austria	Bel-gium	Canada	Den-mark	Finland	France	Ger-many	Nether-lands	Norway	Port-ugal	Spain	Sweden	Switzer-land	UK
OLS estimates														
Constant	0.70	0.00	0.05	0.15	0.40	0.27	0.86	-0.09	0.52	0.95	0.13	0.46	0.35	0.06
	(0.45)	(0.64)	(0.81)	(0.43)	(0.73)	(0.21)	(1.04)	(0.56)	(0.51)	(1.09)	(0.80)	(0.60)	(1.07)	(0.50)
	[1.54]	[0.00]	[0.06]	[0.34]	[0.54]	[1.28]	[0.82]	[-0.17]	[1.02]	[0.87]	[0.17]	[0.77]	[0.32]	[0.13]
Lagged inflation	0.22	0.15	0.24	0.37	0.17	0.14	0.26	0.30	0.23	-0.10	-0.12	0.28	0.14	0.26
	(0.12)	(0.12)	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.11)
	[1.89]	[1.27]	[2.09]*	[3.29]**	[1.46]	[1.18]	[2.21]*	[2.62]**	[1.98]*	[-0.81]	[-0.98]	[2.38]*	[1.14]	[2.25]*
Observations:	71	71	71	71	71	71	71	71	71	71	71	71	71	71
R-squared:	0.05	0.02	0.06	0.14	0.03	0.02	0.07	0.09	0.05	0.01	0.01	0.08	0.02	0
Instrumental variables estimates														
Constant	0.86	-0.05	0.05	0.10	0.07	0.21	0.74	-0.08	0.28	0.65	0.00	0.24	0.46	0.07
	(0.51)	(0.68)	(0.81)	(0.44)	(0.99)	(0.23)	(1.06)	(0.56)	(0.56)	(1.21)	(1.03)	(0.64)	(1.10)	(0.50)
	[1.68]	[-0.07]	[0.06]	[0.24]	[0.08]	[0.92]	[0.70]	[-0.15]	[0.50]	[0.54]	[0.00]	[0.38]	[0.42]	[0.13]
Lagged inflation	0.04	0.54	0.25	0.53	1.03	0.32	0.35	0.39	0.66	0.34	0.67	0.63	-0.09	0.31
	(0.27)	(0.22)	(0.24)	(0.14)	(0.39)	(0.21)	(0.15)	(0.15)	(0.17)	(0.32)	(0.32)	(0.20)	(0.17)	(0.15)
	[0.14]	[2.50]*	[1.04]	[3.74]**	[2.67]**	[1.51]	[2.26]*	[2.65]**	[3.88]**	[1.07]	[2.07]*	[3.15]**	[-0.53]	[2.15]*
Observations:	71	71	71	71	71	71	71	71	71	71	71	71	71	71
Instrumental variables diagnostic tests														
First-stage F-statistic	16.19	37.32	23.57	121.31	14.34	33.07	90.97	104.76	91.79	14.41	19.56	42.17	67.16	129.45
DHW tests *p*-value	0.45	0.01	0.97	0.05	0.00	0.29	0.36	0.31	0.00	0.10	0.00	0.01	0.06	0.55

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.

Table 6: Comparison of AR parameter estimates using OLS and IV on US and international data, 1843-1913

	US data, instrument: series-specific median	US data, instrument: commodity price inflation*	US data, instrument: Canadian inflation*	International data (average 14 countries)	Average
OLS estimate	0.44	0.44	0.24	0.18	0.33
IV estimate	0.69	0.93	0.77	0.44	0.71
Ratio OLS/IV estimate	156.82	211.36	320.83	244.44	233.37

Note: *IV estimate averages exclude Hansen series for which the parameter estimate is in excess of 1. Including these, the averages are 0.97 and 0.92 for the estimates using commodity price inflation and Canadian inflation as instruments, respectively.

Figure 1: US inflation series, 1842-1913

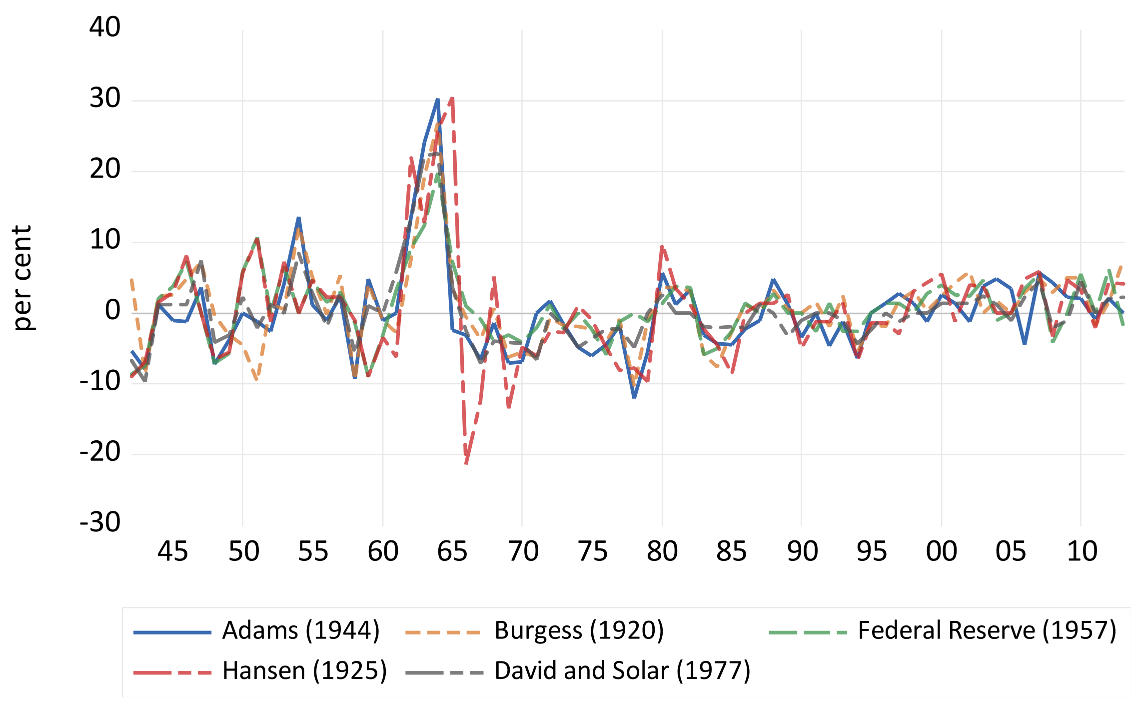


Figure 2: Median and range of inflation across 14 countries, 1842-1913

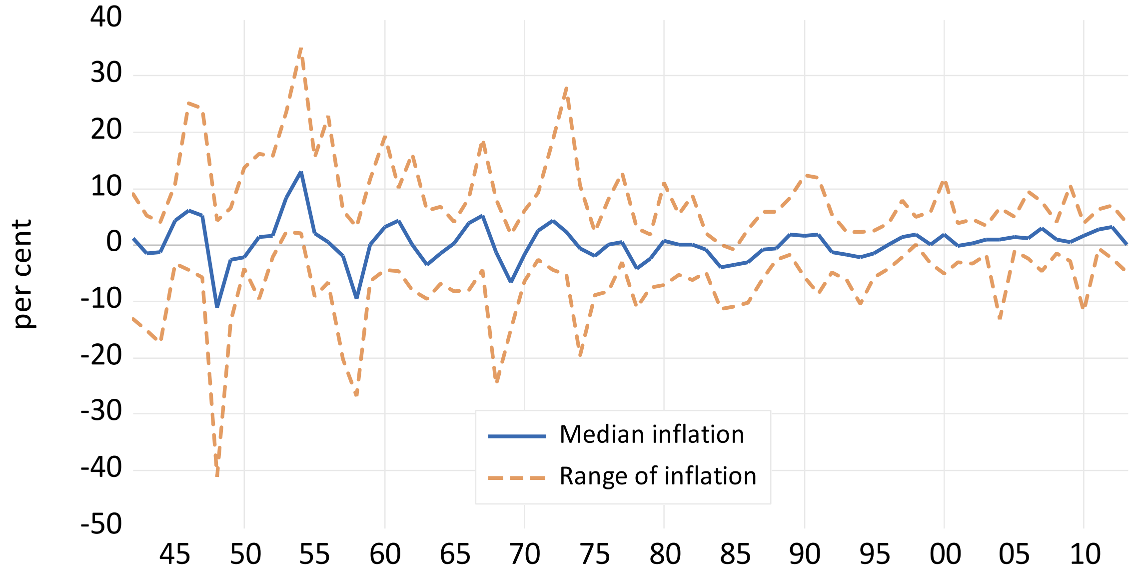
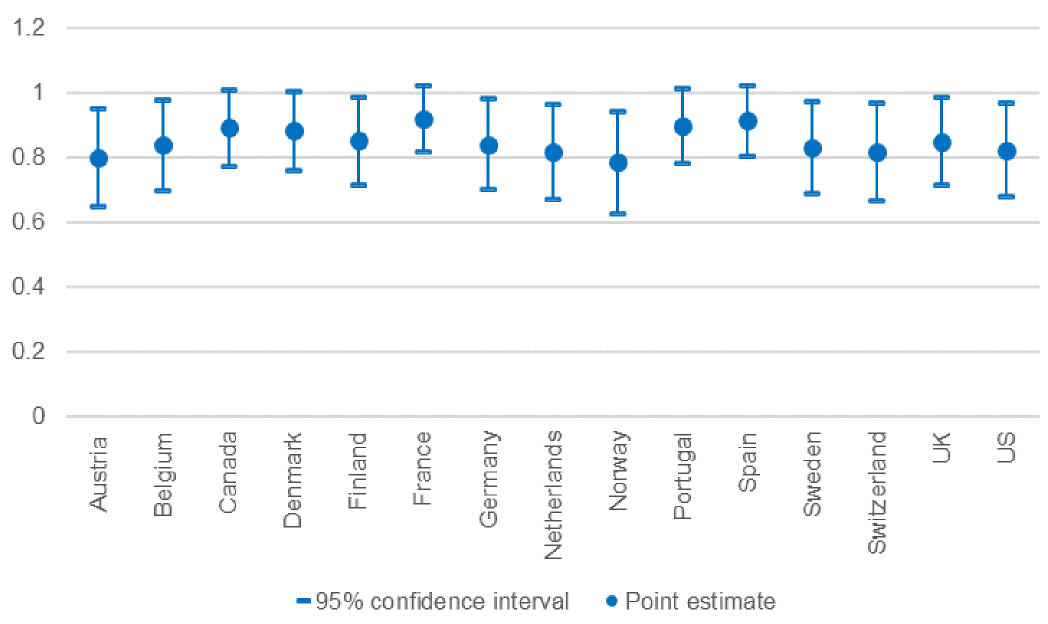


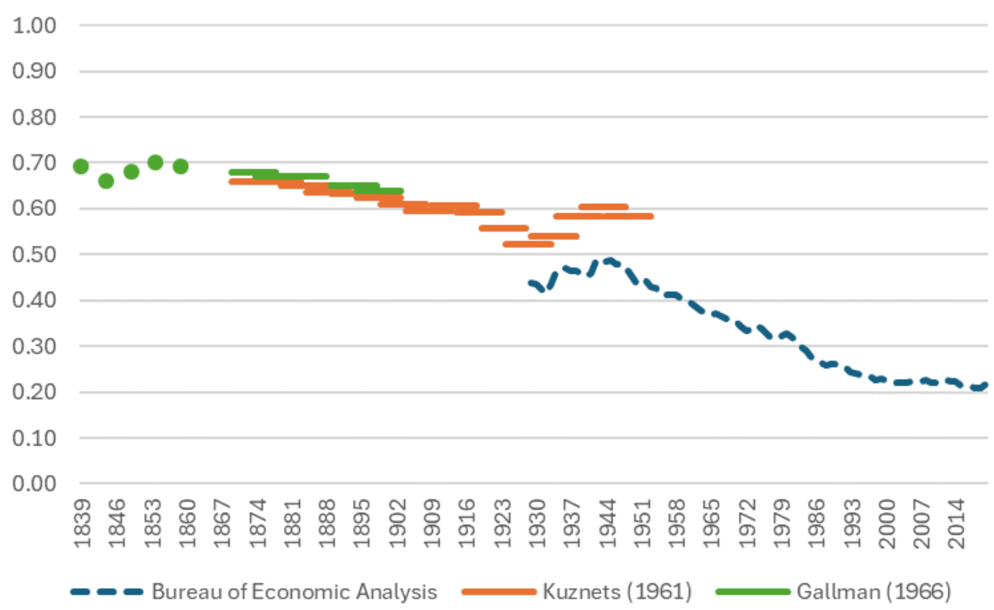
Figure 3: Estimated AR(1) parameter ($\boldsymbol{\beta_i}$), 15 countries, 1960-2020, from inflation equation:

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$



Sources: See Appendix C.

Figure 4: Share of non-durables in personal consumption expenditures, 1839-2020



Sources: BEA Table 2.3.5. Personal Consumption Expenditures by Major Type of Product, Kuznets (1961) Table R-13, p. 522, Gallman (1961) Table 5, p.18.

Notes: Shares calculated in current prices. Data from Kuznets (1961) and data from 1869-1903 from Gallman (1961) are overlapping decadal averages. Non-durables consumption expenditures from Kuznets (1961) and Gallman (1966) are calculated as the sum of expenditures on "perishables" and "semi-durables".

Appendix A: US Data

Components of composite indices, 1850-1913.

David and Solar. 1842-51: David and Solar, 1851-1860: Hoover (1960), 1860-1880: Lebergott (1964), 1880-1890: Long (1960), 1890-1914: Rees (1961).

Federal Reserve. 1842-1860: US Senate (1893)[35], 1860-1880: Mitchell (1908), 1880-1890: Burgess (1920), 1890-1910: Douglas (1926), 1910-1913; Commission on the Necessaries of Life (1920).

Hansen. 1842-1890: US Senate (1893), 1890-1913: Hansen (1925).

Instrumental variables: calculation of the series-specific medians, 1842-1913.

Instruments are calculated as the median of available series at each point in time. Below, we list all series used in the calculation of instruments. In all cases, the median is calculated using between three and five series. Because some of the series are composites, there are a few years when the same underlying series is used to construct more than one of our five series. Specifically, the Federal Reserve and Hansen (1925) use the same underlying series for the period 1842-1860 and the Federal Reserve also uses Burgess (1920) for the 10 years from 1880-1890. To avoid creating spurious correlation with the instrument, in these cases, we drop the overlapping observations from the series used to construct the instrument.

Finally, to reduce the noise in the median as much as possible, we also include in the median calculations data from two additional series (Hoover (1960) and Douglas (1926)). Subsamples of both series are used in the composite series used in the analysis, but some observations are not included in any of our series of interest. We therefore include these additional data points in our median calculations.

The series used to instrument each series are as follows:

Instrument for the Adams series. (1944). David and Solar (1842-1913), Hansen (1842-1914), Federal Reserve (1860-1880, 1890-1913), Hoover (1960) (1860-1890), Douglas (1926) (1910-1913), Burgess (1920) (1842-1913).

[35] Also referred to as the Aldrich Report. The specific series is referred to as Falkner-2 by Officer (2014).

Instrument for the Burgess series. David and Solar (1842-1913), Hansen (1842-1914), Federal Reserve (1860-1880, 1890-1913), Hoover (1960) (1860-1890), Douglas (1926) (1910-1913), Adams (1944) (1842-1913).

Instrument for the Federal reserve series. David and Solar (1842-1913), Hansen (1860-1913), Hoover (1960), (1860-1890), Douglas (1926) (1910-1913), Adams (1944) (1842-1913), Burgess (1920) (1842-1880, 1890-1913).

Instrument for the Hansen series. David and Solar (1842-1913), Federal Reserve (1860-1880, 1890-1913), Hoover (1960) (1860-1890), Douglas (1926) (1910-1913), Adams (1944) (1842-1913), Burgess (1920) (1842-1913).

Instrument for the David and Solar series. Hansen (1842-1914), Federal Reserve (1860-1880, 1890-1913), Hoover (1960) (1860-1890), Douglas (1926) (1910-1913), Adams (1944) (1842-1913), Burgess (1920) (1842-1913).

Appendix B: Sources for international data

Historical international data sources

Country	Source
Austria	Mühlpeck, Sandgruber and Woitek (1979)
Belgium	Mitchell (2003)
Canada	1851-1870: Geloso (2019, Table A4); 1871-1900; Geloso and Hinton (2020); 1901-1910: Historical Statistics of Canada, Series K33, https://www150.statcan.gc.ca/n1 /pub/11-516-x/sectionk/4057753-eng.htm; 1911-1913: Bertram and Percy (1979).
Denmark	Abildgren (2009)
Finland	Heikkinen (1997)
France	Mitchell (2003)
Germany	Mitchell (2003)
Netherlands	Arthur van Riel, http://iisg.nl/hpw/brannex.php
Norway	Grytten (2004)
Portugal	Instituto Nacional de Estatística (2001), Table 8.1
Spain	Barquin (2001), pp. 61-73 and 212-15
Sweden	Edvinsson and Söderberg (2010)
Switzerland	Historical Statistics of Switzerland (2012)
UK	FRED, fred.stlouisfed.org

Recent international data sources, including US (used in Figure 3):

OECD databank, annual data from 1960 to 2020.

Appendix C: Estimation using Seemingly Unrelated Regression (SUR)

The main analysis in the paper is carried out using OLS and instrumental variables (two-stage least squares, TSLS). In this Appendix, we carry out the same regression analysis using SUR and instrumental variables (in the case of SUR, this involves using three-stage least squares, 3SLS). We do not report these in the main text as, if the model were mis-specified, the 3SLS results would be misleading. The equations estimated and instruments used are the same as in the main text.

C1 US results using the series-specific median as an instrument

The results are presented in Table C1. The AR parameter estimates using SUR are significant at the 5% level in four of the five cases, the exception being when the Hansen (1925) series is used. The average AR parameter estimate is 0.22, well below the average parameter when using OLS (0.44). The 3SLS estimates are statistically significant in all cases, and the average AR parameter estimate is 0.69.[36] DHW tests reject the null in 3 cases (the exceptions, as in the OLS analysis, are Adams (1944) and David and Solar (1977) which return p -values of 0.10 and 0.13, respectively), however, a joint test returns a p -value of 0.00. Overall, these results are similar to when OLS is used, and consistent with measurement error being present.

C2 US results using commodity price inflation as an instrument

The 3SLS estimates using commodity price inflation as the instrument are presented in the lower Panel of Table C1. The average of the estimated AR parameters is 0.97, and the DHW tests reject the null in all cases with p-values of 0.00.

C3 US results using Canadian inflation as an instrument

The results are presented in Table C2. As with the analysis in the main text, the estimates here exclude the Civil War period and do not include the regression for the Hansen (1925) series since the first stage F-statistic is less than 10. Using SUR, the average of the AR parameters is 0.25, and the parameters are significant at the 5% level in the case of the Federal Reserve (1957) and David and Solar (1977) and at the 10% level for the other two.

---

[36] The point estimates for TSLS and 3SLS are identical when all equations in the system are just identified, as is the case here (Kapteyn and Fiebig (1981)).

Using 3SLS, the AR parameter is statistically significant at the 5% level in all cases, averaging 0.89. DHW tests indicate that the null can be rejected at the 5% level in three cases and at the 10% level in the fourth (Federal Reserve (1957)). A joint test returns a p -value of 0.00.

Thus, in the case of all three instruments used with the US data, the increase in the parameter estimate using 3SLS compared to SUR is greater than when comparing TSLS to OLS. In both cases, the increases in parameter estimates using IV are statistically significant. Overall, this reinforces the results in the main text.

C4 International data results using the country-specific median as an instrument

The results for the international data are presented in Table C3. The average AR parameter estimate using SUR is just 0.05 and the parameter is significant in just five of 14 cases. Using 3SLS to estimate the system, the average parameter value is 0.43 and it is significant at the 5% level in nine cases. DHW tests indicate that the null can be rejected in six cases. A further three cases are significant at the 10% level (Netherlands, Portugal and the UK). A joint DHW test returns a p -value of 0.00. Overall, these results are in line with those in the main text and are supportive of the hypothesis of measurement error being present.

Table C1: SUR and 3SLS estimates of inflation equation, 1843-1913

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Five US inflation series, cross-sectional median as the instrument

Dependent Variable	**Adams (1944)**	**Burgess (1920)**	**Federal Reserve (1957)**	**Hansen (1925)**	**David and Solar (1977)**
			OLS results		
Constant	0.14 (0.68) [0.20]	0.55 (0.63) [0.87]	0.67 (0.49) [1.37]	0.40 (0.90) [0.44]	0.23 (0.49) [0.47]
Lagged inflation	0.18 (0.07) [2.40]*	0.28 (0.08) [3.67]**	0.23 (0.09) [2.65]**	0.11 (0.10) [1.10]	0.32 (0.07) [4.79]**
Observations:	71	71	71	71	71
R-squared:	0.12	0.15	0.17	0.05	0.29
	(a) 3-stage least squares results using series-specific median as an instrument				
Constant	0.11 (0.69) [0.16]	0.27 (0.66) [0.40]	0.26 (0.53) [0.50]	0.20 (1.10) [0.18]	0.19 (0.49) [0.39]
Lagged inflation	0.50 (0.12) [4.17]**	0.68 (0.13) [5.10]**	0.77 (0.15) [4.99]**	0.93 (0.21) [4.33]**	0.57 (0.10) [5.67]**
Observations:	71	71	71	71	71
DHW test p - value	0.10	0.01	0.00	0.00	0.13
	(b) 3-stage least squares results using commodity price inflation as an instrument				
Constant	0.09 (0.75) [0.12]	0.04 (0.78) [0.05]	0.11 (0.60) [0.19]	0.15 (1.25) [0.12]	0.14 (0.53) [0.26]
Lagged inflation	0.81 (0.16) [5.16]**	1.00 (0.23) [4.39] **	0.98 (0.22) [4.42] **	1.16 (0.32) [3.61] **	0.91 (0.15) [5.93] **
Observations:	71	71	71	71	71
DHW test p - value	0.00	0.00	0.00	0.00	0.00

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.

Table C2: SUR and 3SLS estimates of inflation equation, 1843-1860, 1966-1913

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Four US inflation series, Canadian inflation as the instrument

Dependent Variable	**Adams (1944)**	**Burgess (1920)**	**Federal Reserve (1957)**	**David and Solar (1977)**
		OLS results		
Constant	-0.62	-0.01	0.07	-0.52
	(0.53)	(0.54)	(0.45)	(0.39)
	[-1.16]	[-0.02]	[0.16]	[-1.33]
Lagged inflation	0.22	0.22	0.30	0.26
	(0.12)	(0.12)	(0.11)	(0.12)
	[1.86]	[1.82]	[2.67]**	[2.21]*
Observations:	66	66	66	66
R-squared:	0.05	0.05	0.10	0.07
		3-stage least squares results		
Constant	-0.19	-0.03	-0.01	-0.04
	(0.61)	(0.78)	(0.51)	(0.49)
	[-0.31]	[-0.03]	[-0.03]	[-0.07]
Lagged inflation	0.69	1.25	0.77	0.85
	(0.22)	(0.44)	(0.32)	(0.25)
	[3.11]**	[2.85]**	[2.42]*	[3.35]**
Observations:	66	66	66	66
		DHW test results		
p -value	0.00	0.00	0.06	0.00

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively. The Hansen (1925) series is not included in the analysis since the first-stage F-statistics is less than 10.

Table C3:
SUR and 3SLS estimates, 14 economies, 1843-1913, of inflation equation

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Dependent variable:	**Austria**	**Bel-gium**	**Canada**	**Den-mark**	**Finland**	**France**	**Ger-many**	**Nether-lands**	**Norway**	**Port-ugal**	**Spain**	**Sweden**	**Switzer-land**	**UK**
							OLS estimates							
Constant	0.79	0.02	0.05	0.19	0.43	0.34	1.07	-0.11	0.56	1.06	0.12	0.62	0.45	0.06
	(0.44)	(0.63)	(0.80)	(0.42)	(0.72)	(0.21)	(1.02)	(0.55)	(0.50)	(1.07)	(0.79)	(0.59)	(1.06)	(0.50)
	[1.78]	[0.03]	[0.06]	[0.45]	[0.60]	[1.65]	[1.04]	[-0.20]	[1.13]	[0.99]	[0.15]	[1.05]	[0.42]	[0.12]
Lagged inflation	0.12	-0.05	0.24	0.18	0.07	-0.06	0.10	0.14	0.15	-0.26	-0.04	0.02	-0.06	0.07
	(0.09)	(0.09)	(0.09)	(0.06)	(0.09)	(0.09)	(0.08)	(0.06)	(0.07)	(0.09)	(0.10)	(0.08)	(0.07)	(0.06)
	[1.33]	[-0.51]	[2.79]**	[3.05]**	[0.83]	[-0.64]	[1.32]	[2.34]*	[2.16]*	[-3.03]**	[-0.40]	[0.21]	[-0.86]	[1.11]
Observations:	71	71	71	71	71	71	71	71	71	71	71	71	71	71
R-squared:	0.04	-0.02	0.06	0.10	0.02	-0.02	0.04	0.07	0.05	-0.02	0.01	0.01	-0.02	0.03
							Two-stage least squares estimates							
Constant	0.86	-0.05	0.05	0.10	0.07	0.21	0.74	-0.08	0.28	0.65	0.00	0.24	0.46	0.07
	(0.51)	(0.67)	(0.80)	(0.43)	(0.97)	(0.22)	(1.04)	(0.55)	(0.55)	(1.19)	(1.01)	(0.64)	(1.09)	(0.50)
	[1.71]	[-0.07]	[0.06]	[0.24]	[0.08]	[0.93]	[0.71]	[-0.15]	[0.51]	[0.55]	[0.00]	[0.39]	[0.42]	[0.13]
Lagged inflation	0.04	0.54	0.25	0.53	1.03	0.32	0.35	0.39	0.66	0.34	0.67	0.63	-0.09	0.31
	(0.27)	(0.21)	(0.23)	(0.14)	(0.38)	(0.21)	(0.15)	(0.15)	(0.17)	(0.32)	(0.32)	(0.20)	(0.17)	(0.14)
	[0.14]	[2.54]*	[1.06]	[3.79]**	[2.71]**	[1.53]	[2.29]*	[2.69]**	[3.94]**	[1.08]	[2.10]*	[3.19]**	[-0.54]	[2.18]*
Observations:	71	71	71	71	71	71	71	71	71	71	71	71	71	71
							DHW test results							
p-value	0.46	0.00	0.69	0.11	0.00	0.14	0.49	0.08	0.00	0.06	0.00	0.00	0.04	0.07

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.

Appendix D: Results for the extended sample period, 1800-1913

Since instrumental variable estimates are valid asymptotically, as a robustness, we re-estimate the analysis in the main text using the long sample period spanning the years 1802 1913. We do so for the two US series that are available for this sample period (Adams (1944) and David and Solar (1977) using Canadian inflation as the instrument (we do not use the median as an instrument as we have just two series) and for the 10 countries for which we have data extending back to 1800. This gives us 107 observations in the case of the US and 112 in the case of the international data (the difference in length arises from dropping the Civil War), or at least a 50% increase in observations compared to the analysis in the main text.

The results for the US series are presented in Table D1. The estimated AR parameters for both series are very similar when estimated using OLS, at 0.14 on average. Using Canadian inflation as the instrument, the first-stage F-statistics are both in excess of 10 and the parameter estimates using IV more than double, to 0.50 for the Adams series, and 0.60 for the David and Solar series. DHW tests indicate that the IV estimates are statistically different from the OLS estimates.

The results for the international data are included in Table D2. While data are available for ten countries (Austria, Canada, Denmark, Finland, Netherlands, Norway, Sweden, Switzerland and the UK) for this sample period, our first stage F-statistics are less than 10 in the case of four of these countries (Canada, Denmark, Finland and Norway). We therefore proceed with the analysis for the remaining six countries. The OLS estimates of the AR parameter average 0.21 for the six countries and the parameter is statistically significant at the 5% level in three cases (Netherlands, Sweden and UK). The IV estimates of the AR parameter average 0.45 across the six countries and are statistically significant in the same three countries. DHW tests indicate that the IV results are statistically different from the OLS estimates in three cases (Netherlands, Sweden and UK again), and a joint test rejects the null (*p*-value = 0.00).

Overall, the results over a longer sample period are consistent with measurement error being present in the data and support the findings in the main text.

Table D1: OLS and IV estimates of inflation equation, 1800-1913

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Two US inflation series, Canadian inflation as the instrument

Dependent Variable	**Adams (1944)**	**David and Solar (1977)**
	OLS results	
Constant	-0.85 (0.49) [-1.74]	-0.75 (0.46) [-1.63]
Lagged inflation	0.15 (0.10) [1.53]	0.13 (0.10) [1.35]
Observations:	107	107
R-squared:	0.02	0.02
	Instrumental variables results	
Constant	-0.50 (0.55) [-0.91]	-0.36 (0.54) [-0.67]
Lagged inflation	0.50 (0.19) [2.55] *	0.60 (0.23) [2.65]**
Observations:	107	107
	Instrumental variables diagnostic tests	
First-stage F-statistic	40.09	31.70
DHW test *p*-value	0.02	0.01

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.

Table D2: OLS and IV estimates, 6 economies, 1800-1913, of inflation equation

$$\pi_{i,t} = \alpha_i + \beta_i \pi_{i,t-1} + \zeta_{i,t}$$

Dependent Variable	**Austria**	**Netherlands**	**Portugal**	**Sweden**	**Switzerland**	**UK**
			OLS results			
Constant	0.42	-0.28	0.04	0.91	-0.07	-0.15
	(0.54)	(0.47)	(1.03)	(0.58)	(0.87)	(0.51)
	[0.78]	[-0.60]	[0.04]	[1.57]	[-0.08]	[-0.30]
Lagged inflation	0.04	0.25	0.09	0.20	0.11	0.26
	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
	[0.46]	[2.68]**	[0.98]	[2.15]*	[1.17]	[2.80]**
Observations:	112	112	112	112	112	112
1R-squared:	0.00	0.06	0.01	0.04	0.01	0.07
			Instrumental variables results			
Constant	0.34	-0.17	0.05	0.16	-0.04	-0.06
	(0.56)	(0.49)	(1.04)	(0.77)	(0.87)	(0.53)
	[0.60]	[-0.36]	[0.05]	[0.20]	[-0.05]	[-0.12]
Lagged inflation	0.25	0.53	0.25	0.89	0.19	0.57
	(0.28)	(0.15)	(0.29)	(0.30)	(0.16)	(0.17)
	[0.88]	[3.54]**	[0.86]	[2.99]**	[1.23]	[3.34]**
Observations:	112	112	112	112	112	112
			Instrumental variables diagnostic tests			
First-stage F-statistic	15.19	79.63	13.58	18.57	59.47	50.35
DHW test *p*-value	0.43	0.01	0.57	0.00	0.50	0.02

Notes: Standard errors in parenthesis, t-statistics in brackets. */** denotes significance and the 5% and 1% level respectively.