DISCUSSION PAPER SERIES

No. 9504

INTEGRATION IN THE ENGLISH WHEAT MARKET 1770-1820

Liam Brunt and Edmund Cannon

ECONOMIC HISTORY and INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP9504.asp

ISSN 0265-8003

INTEGRATION IN THE ENGLISH WHEAT MARKET 1770-1820

Liam Brunt, The Norwegian School of Economics and Business Administration and CEPR
Edmund Cannon, University of Bristol

Discussion Paper No. 9504
June 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9504

June 2013

ABSTRACT

Integration in the English wheat market 1770-1820*

Cointegration analysis has been used widely to quantify market integration through price arbitrage. We show that total price variability can be decomposed into: (i) magnitude of price shocks; (ii) correlation of price shocks; (iii) between-period arbitrage. All three measures depend upon data frequency, but between-period arbitrage is most affected. We measure variation of these components across time and space using English weekly wheat price data, 1770-1820. We show that conclusions about arbitrage are sensitive to the precise form of cointegration model used; different components behave differently; and different factors – in terms of transport and information – explain behaviour of different components. Previous analyses should be interpreted with caution.

JEL Classification: N73, Q11 and R41
Keywords: domestic trade, economic integration, England and Wales, grain markets, time-series cointegration and transport

Liam Brunt
Department of Economics
Norwegian School of Economics and Business Administration
Helleveien 30
5045-Bergen
NORWAY

Email: liam.brunt@nhh.no

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165274

Edmund Cannon
University of Bristol
Department of Economics
8 Woodland Road
Bristol BS8 1TN

Email: edmund.cannon@bristol.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154984

Submitted 04 June 2013

1. Introduction

The last decade has seen many papers analysing market integration using data sets for many countries and time periods and a variety of econometric methods: Federico (2012) provides a comprehensive survey. The simplest way to analyse integration is to look at a measure of contemporaneous price dispersion (typically the coefficient of variation) and then see how this changes over time (see, for example, Jacks, 2011; he uses the same data that we use in this paper). Alternatively one can use more sophisticated econometric techniques to analyse the price series jointly (i.e. a vector-auto-regressive approach such as that of Ejrnæs, Persson and Rich, 2008, or Studer, 2008; these build on the seminal work of Ravallion, 1986).

Federico (2012) notes that there are two requirements for distinct markets to be integrated: (i) the long-run equilibrium should have similar prices in the different markets (the Law of One Price, or "LOOP"); and (ii) in the short run price differences should correct relatively quickly if the equilibrium is disturbed (which Federico refers to as "efficiency"). Econometric time-series models using cointegration techniques potentially have the ability to confirm that LOOP characterises prices in the long run and to estimate the speed with which price differentials are arbitraged away.

Phillips (1991) shows that estimating and testing long-run relationships – which is LOOP for our purposes – is typically unaffected by data problems, so long as there is a sufficiently long span of data. Since many economic historical studies use very long spans of data, this is generally not a problematic aspect of the literature. Therefore most of our analysis here is devoted to the second issue – namely the speed of adjustment towards equilibrium – and we take LOOP as given.

In our analysis we make three methodological points. First, we quantify the significance of using high-frequency (weekly) data to estimate the dynamics of price movements. Although it is well understood that higher frequency data are better, we have seen no attempt to put a number on just how much empirical results depend upon the frequency of observation.[3] Since we have weekly data, we are able to say how our results would look if we used weekly, monthly or annual data.

[3] Taylor (2001) shows that parameter estimates will be inconsistent if the data consist of prices averaged within observation periods (for example, if we use monthly data consisting of the average of daily prices within the month); this phenomenon is called

Second, much of the literature measures market efficiency using the "half life" (that is, the time for half of a price disequilibrium to be corrected). It does this using a Vector Error Correction Model (VECM) with no additional lags. We show that this model typically does not fit the data and hence conclusions about market efficiency using this model are misleading. In particular, it is not possible to construct a simple estimate of the half life and it is better to estimate the full impulse response function.

Third, we show how the variance of prices can be decomposed into the disturbances causing prices to move apart, and the process by which they move back together. This goes some way to answering the critique of Federico (2010, 2012), who suggests that market integration should be measured by sigma convergence.

Stepping back from these econometric concerns, we turn to the determinants of market efficiency during the Revolutionary and Napoleonic Wars. The period for which we have data (1770-1820) coincides with the large expansion of the canal network (Priestley, 1831) and the improvement of the road network (Albert, 1972; Pawson, 1977).[4] But what was the effect of better transport on market integration and are we able to measure it?

One way to measure the effect of new transport is the social savings method, as has been used by Bogart (2005b, 2011) for turnpiking (we are unaware of any attempt to do a social savings approach for canals). But the social savings approach is only really useful where one form of transport replaces another, whereas one of the achievements of canals was to allow heavy goods, such as coal, to be transported over distances that were previously impractical (Leunig, 2010). Donaldson (2010) suggests instead using a general equilibrium approach to quantify the effects of a transport innovation on the whole economy and he applies this to Indian railroads in the nineteenth century; this makes a lot of sense in that situation because Indian railroads were not replacing any

---

temporal aggregation and the direction of bias means that the efficiency of the market (i.e. estimated speed of convergence to equilibrium) will be under-estimated. The data that we shall use in this paper are weekly data and in many instances markets only traded once or twice per week, so temporal aggregation is not our primary concern.

[4] Bagwell (1974) and Chartres and Turnbull (1983) provide evidence that the volume of road transport increased considerably over this period. Timmins (2005) and Gerhold (1993, 1996) show that the increase in road mileage was accompanied by better road and transport technologies respectively.

existing transport technology. Identifying the causal effect of transport is also difficult because causation runs in several directions: it may be easier to improve transport during times of greater economic activity (Ward, 1974), in which case changes in transport and market structure may be coincident but not causally related. In addition, poor market integration is likely to be a key incentive to improve transport, in which case the researcher can observe the effect of transport on market integration only in a non-random sub-set of markets. Thus, using instrumental variable techniques, Donaldson and Hornbeck (2012) obtain larger estimates for the effect of railways in America than the original estimates of Fogel (1964).

Bogart (2005a, 2011) argues that the improvements to both roads and canals were due to institutional changes that made these forms of investment easier, suggesting that improved transport was a cause of better market integration. However, such institutional changes only enabled improvements in general; they do not explain the structure or sequencing of specific canal and road improvement. Hence this does not negate the criticism that the transport improvements that we see are non-random.

From the perspective of the grain market, however, it is likely that endogeneity of improvements in transport may be relatively unimportant. Most canals were built with the primary purpose of transporting coal or manufactures and so the benefits to the grain market would have been a side effect; evidence on the use of canals to transport different commodities during this period can be found in Maw (2009). Similarly, although Gerhold (1996) argues that short-distance transport of grain on roads was an area where productivity growth was high, it is difficult to imagine that this alone would have provided an incentive to improve roads that were used for a variety of other goods.

Furthermore, better integration of grain markets was probably due less to improved productivity in freight than in passenger travel, because this is the speed at which information and news is transmitted. Since grain was stored week-to-week, temporal arbitrage would mean that news about other markets would affect any given market when the news arrived, rather than the point at which any grain arrived. For this reason we view better roads as potentially having a large effect on high-frequency movements in prices.

To test the importance of information transmission, we estimate models of market efficiency with both transport and communication variables as possible explanators. We show that market efficiency increased in England during the period 1770-1820, even during the Revolutionary and Napoleonic Wars. During that conflict the magnitude of shocks to the grain market increased, but there was an underlying improvement in market efficiency due to improved roads, canal building and increased newspaper circulation.

Ironically, none of the transport or communication variables were responsible for shortening the measure most widely used in the economic history literature – namely, the half life. Half lives did shorten throughout the period 1770-1820, and even fell during the period of the Revolutionary and Napoleonic Wars. However, we are unable to find any variables that correlate with our half life estimates. Instead we find that, conditional on the general state of the economy, improved transport and communication resulted in *smaller shocks to prices* and also prices *moving more closely together*. These are the routes by which better transport and communication generated arbitrage. We thus make a methodological contribution to the literature and illustrate it with the important case of England during the Industrial Revolution.

The rest of our paper is organised as follows. In section 2 we describe our data quite carefully and summarise it with a series of measures that have been used elsewhere in the literature. This approach suggests that very little changed in the economy. But such an inference would be incorrect: several important changes occurred, which happened to cancel each other out (better transport increased market efficiency, while the Revolutionary and Napoleonic Wars increased market turbulence). In section 3 we describe the econometric issues formally. Section 4 presents examples of our econometric procedures together with our estimates of half lives and other measures of market efficiency. In section 5 we correlate our measures of market efficiency with transport and communication variables. Section 6 concludes. blah

2. Wheat Prices 1770-1820

In this section we provide an overview of the English wheat market in the period 1770 to 1820: our data are described in detail in Brunt and Cannon (2013). During this period prices were collected weekly for a large number of towns in England and Wales: these town-level prices were averaged for each county and then published about a week later in the official publication the *London Gazette*. This means that we have weekly data for the average price for forty counties (39 English counties and one Welsh).[5] We are

[5] Data for London are available only until 1794 and we do not use them because the London market is not a representative area. Monmouthshire was treated as an English

unable to extract the town-level prices from the available data as the original town level prices were never published and have presumably been lost. There is good reason for confining our attention to wheat prices. First, the markets for oats, beans and peas have more missing observations. Second, the quantities traded for those crops between 1770 and 1820 were probably relatively unimportant (data on the quantities traded are not available for this earlier period, so we cannot test this hypothesis directly, but we know that the quantities were relatively small from 1820 onwards). Third, there are even more problems with barley prices: although barley sales were relatively large, they were concentrated in a relatively small part of the year (September-November) and the market for the rest of the year is so thin that the prices are unlikely to be informative.

This leaves wheat as our focus of analysis. Given data from 1818 and later periods, we know that this grain was traded steadily throughout the year, as would be expected from the crop that was fairly easy to store and which provided the main foodstuff in the UK at this time (Petersen, 1995). There are also relatively few missing observations: out of 2604 weekly observations many counties are missing only a few data points (the worst county, Hereford, lacks just 61).

Figure 1 about here (wheat price)

Figure 1 illustrates the movement of grain prices over the whole period, plotting the minimum and maximum price in each week. The range of prices in each week is large – on average about 33 pence, or a third of the price.

We now turn to more formal measures of price behaviour. We write the natural logarithm of the price in county *i* in time period *t* as either $p_{i.t}$ or p_t^i depending on which notation is more convenient. When we wish to distinguish annual from weekly data explicitly we shall write $p_{i.y}$ or $p_{i.w}$ respectively and when we wish to refer to all 52 (or 53) weeks within a year we shall write $w \in y$: for example, the average price within a year is of the form

---

county at this time, possibly because it was on an English county circuit. Data for other parts of Wales were originally published in even more aggregate form; when published for individual Welsh counties, there are many missing observations, so we do not use these data.

$$(1) \qquad \tilde{p}_{i.y} \equiv \sum_{w \in y} p_{i.w} \Big/ 52$$

$$\text{e.g. } \tilde{p}_{\text{Bedfords.1781-2}} \equiv \frac{p_{\text{Bedfords.1781,week 45}} + p_{\text{Bedfords.1781,week 46}} + \cdots + p_{\text{Bedfords.1782,week 44}}}{52}$$

Note that we calculate yearly average data using harvest years, which we assume to start in the first week in October (typically the 45th week of the year) and to finish in the last week in September. Henceforth, whenever we refer to annual data we are referring to *harvest* rather than *calendar* years.

We start by measuring the standard deviation in prices. In principle we could calculate this for all 2604 weeks for which we have data, but we summarise our results by reporting annual averages (this also smooths out idiosyncratic changes from week to week). As an informal check on the effect on averaging we calculate both the standard deviation of annual averages and also the annual average of weekly standard deviations, which are defined formally as

$$(2) \qquad \sigma_y^{(1)} \equiv \sqrt{\frac{\sum_i \left(\tilde{p}_{i.y} - \overline{\tilde{p}}_y\right)^2}{40}}; \qquad \sigma_y^{(2)} \equiv \frac{\sum_i \sqrt{\sum_{w \in y} \left(p_{i.w} - \overline{p}_w\right)^2}}{40}$$

and illustrated in Figure 2: the two measures are very similar.[6] The standard deviation in any week is typically about 0.08, which we can interpret by saying that the standard deviation of prices was consistently about 8 per cent of the price.

Figure 2 about here (dispersion of prices between counties)

From both Figures 1 and 2, and the related calculations, there appears to be no systematic change over the fifty year period: the range of prices does not trend down (which we might expect to be a consequence of greater transport links) and there is no obvious increase during the Napoleonic Wars. In fact, the average value of $\sigma_y^{(1)}$ is 8.0 per cent for 1771-1792 and then 7.9 per cent for 1793-1815, while the average value of $\sigma_y^{(2)}$ is 6.9 per cent to 6.6 per cent in the same two sub-samples respectively (neither of these

[6] There is no significant seasonal pattern in the standard deviation, so most of the difference between the two measures in the graph appears to be due to Jensen's inequality. Note that the standard deviation of log prices is almost identical to Federico's (2011) preferred measure of the coefficient of variation in prices: the correlation coefficient between the two measures on weekly data is 0.996.

changes being statistically significant). So, if transport and war effects were important, then they must have cancelled each other on this measure.

Of course the fact that the standard deviation of prices is fairly constant tells us relatively little about how the prices were interacting with each other. Simple correlations of the price series are uninformative because there is considerable variation in prices, and prices move sufficiently closely together that the underlying trend will dwarf any other effects (typical correlations are 0.98-0.99).

A more interesting question is to ask how relative prices changed over time. To do this we take annual cross sections of the average within-harvest-year prices at the beginning of the harvest year and calculate the correlation with the corresponding prices at the beginning of the following year:

(2)
$$\text{corr}\left(p_{i.y}, p_{i.y+1}\right) = \frac{\sum_i \left(p_{i.y} - \overline{p}_y\right)\left(p_{i.y+1} - \overline{p}_{y+1}\right)}{\sqrt{\sum_i \left(p_{i.y+1} - \overline{p}_{y+1}\right)^2 \sum_i \left(p_{i.y} - \overline{p}_y\right)^2}}$$

If the pattern of relative prices in the different counties were to stay the same, then we should expect this statistic to be high. Figure 3 shows the value of the statistic for consecutive pairs of years over the whole period: given the sample size, these correlations are statistically significantly positive if bigger than 0.29. The only years when the pattern of prices changed much from the previous year are 1772, 1779, 1800 and 1808, suggesting that relative prices changed fairly slowly. In further analysis discussed in Appendix 2 we find that the pattern of relative prices remained remarkably stable over the entire period.

Figure 3 about here (year-on-year correlations of cross sectional prices)

Our final summary characterisation of the data is to see how price differences depend upon proximity of counties, which we describe using the conventional Moran's *I* statistic (more sophisticated measures of spatial correlation would yield similar conclusions):

(4)
$$I_w = \frac{40\sum_j \sum_i a_{ij}\left(p_{j.w} - \overline{p}_w\right)\left(p_{i.w} - \overline{p}_w\right)}{\left(\sum_j \sum_i a_{ij}\right)\left(\sum_k \left(p_{k.w} - \overline{p}_w\right)^2\right)}$$

where $a_{ij} = 1$ if counties *i* and *j* are adjacent and zero other wise. This statistic is calculated on the cross-section of prices for each week of the sample and illustrated in

Figure 4. Under the null hypothesis of no spatial correlation, the expected value of this statistic is –0.024: our calculated *I* statistics average 0.41, typically in the range from 0.2 to 0.6, and are almost invariably statistically significant with (standard Normal) Z-statistics averaging 4.34. Yet again, there appears to be no systematic variation over time (there is no seasonal pattern in the Moran statistic: a regression on seasonal dummies yields a test statistic of $F(52, 2550) = 0.72$ [$p = 0.94$]).

Figure 4 about here (Moran's I)

We summarise our analysis so far by noting that, although prices were very different in the various counties for which we have data, these prices all moved closely together over the entire period. Moreover, from an analysis of summary statistics, their behaviour does not appear to have changed much over the period 1771-1820. There was a high degree of spatial correlation, which did not change much, and the relative prices in different counties was also roughly the same at the beginning of the period as it was at the end. Very similar results can be obtained whether using end-of-year prices or within-year averages, so the frequency of measurement is not a major determinant of our conclusions.

This evidence provides only the most minor support for the idea that improved transport significantly affected grain prices. The most important prediction of any model of falling transport costs would be some form of convergence, some change in relative prices or some change in the relevance of distance. From our analysis of summary statistics in this section, none of these things happened. Therefore we can only conclude that we need to model price behaviour of the individual series much more closely. In the following section we consider a framework for discriminating between different determinants of price movements.

3. Cointegrated prices: explanation and example

It is obvious from the previous section that individual price series show both large and persistent variation over time. This is true not just for our data but for many other data sets. Usually one cannot reject the null hypothesis that any given price series has a unit root, meaning that standard statistical theory will not apply to some estimation and testing procedures. It is also common for price series to move closely together, so that the difference in – or the ratio of – two price series is much less variable and more persistent. This suggests that there is a simple equilibrium relationship between the two price series; if the difference or ratio of prices does not have a unit root then the series are cointegrated. A good introduction to this approach is provided by Ejrnæs (1999) and we shall build on that analysis here.

We consider a fairly general error-correction model (ECM), which illustrates the main points of this paper and can encompass many, but not all, of the other issues that may be relevant. Our starting point is the Data Generating Process (DGP):

$$
(5)\quad
\begin{aligned}
&\begin{bmatrix}\Delta p_t^i \\ \Delta p_t^j\end{bmatrix} = \begin{bmatrix}\alpha_i \\ \alpha_j\end{bmatrix}\left(p_{t-1}^i - p_{t-1}^j\right) + \begin{bmatrix}\mu_i \\ \mu_j\end{bmatrix} + \begin{bmatrix}\pi_{i.i}^{(1)} & \pi_{i.j}^{(1)} \\ \pi_{j.i}^{(1)} & \pi_{j.j}^{(1)}\end{bmatrix}\begin{bmatrix}\Delta p_{t-1}^i \\ \Delta p_{t-1}^j\end{bmatrix} + \cdots + \begin{bmatrix}\pi_{i.i}^{(K)} & \pi_{i.j}^{(K)} \\ \pi_{j.i}^{(K)} & \pi_{j.j}^{(K)}\end{bmatrix}\begin{bmatrix}\Delta p_{t-K}^i \\ \Delta p_{t-K}^j\end{bmatrix} + \begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix}; \\
&\begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix} \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\omega_{i.i} & \omega_{i.j} \\ \omega_{i.j} & \omega_{j.j}\end{bmatrix}\right); \qquad \alpha_i \le 0; \quad \alpha_j \ge 0; \quad \alpha_j - \alpha_i < 1
\end{aligned}
$$

where $\Delta p_t^i \equiv p_t^i - p_{t-1}^i$, which we refer to as the price change.[7] The cointegration equation can be written more compactly in vector notation as

$$
(6)\qquad \Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\gamma}\mathbf{p}_{t-1} + \boldsymbol{\mu} + \sum_{k=1}^{K} \boldsymbol{\pi}^{(k)} \Delta \mathbf{p}_{t-k} + \boldsymbol{\varepsilon}_t; \quad \boldsymbol{\varepsilon}_t \sim N\left(\mathbf{0}, \boldsymbol{\Omega}\right); \quad \boldsymbol{\gamma} \equiv \begin{bmatrix}1 & -1\end{bmatrix}
$$

This simple ECM means that prices change to reduce the disequilibrium regardless of how far apart they are. An alternative to this is a Threshold ECM, where prices adjust if the disequilibrium is large, but not if it is small. A natural interpretation of this model is that arbitrage takes place if the price gap exceeds transport costs, but that a small price disequilibrium does not create profitable arbitrage opportunities and hence no correction occurs. The point at which price behaviour changes can be based on actual transport cost data (e.g. Ejrnæs, Persson and Rich, 2008) or estimated using maximum likelihood (e.g. Ejrnæs and Persson, 2000).

---

[7] To avoid ambiguity, we avoid using the phrase *price difference*, which might imply either the price change $\Delta p_t^i \equiv p_t^i - p_{t-1}^i$ or the market disequilibrium (or price gap) $\boldsymbol{\gamma}\mathbf{p}_t \equiv p_t^i - p_t^j$. Throughout this paper we impose homogeneity in prices, unlike papers such as Sharpe and Weisdorf (2013) who estimate a relationship of the form $p_t^i = \theta p_t^j$ and also include a deterministic trend.

Our data are poorly suited to the application of a TECM, because we do not observe prices at nodal points (i.e. markets within a town) but averages for a small region (i.e. an English county). It is impossible to define a transport cost between two regions because there would be many possible points in region 1 to join to many possible points in region 2; nor is it obvious how the relationship between two regional averages would depend upon the transport costs anyway. For example, even if the gap between the two regions' average prices were less than the average transport cost between the two regions, the gap between the prices of one particular town in the first region and another particular town in the second region could still be further apart than the transport cost between those two towns. In this case there might continue to be arbitrage between those two towns in which case those two prices would continue to move to equality: in which case the averages would also continue to adjust to equilibrium. For this reason we prefer to use the ECM for this particular data set, noting that many of the issues that we raise would also be relevant for a TECM.

If the lagged price changes are not needed then the DGP can be re-written:[8]

$$(7) \qquad \left(\boldsymbol{\gamma}\mathbf{p}_t - \lambda\right) = \left(\boldsymbol{\gamma}\boldsymbol{\alpha} + 1\right)\left(\boldsymbol{\gamma}\mathbf{p}_{t-1} - \lambda\right) + \boldsymbol{\gamma}\boldsymbol{\varepsilon}_t \qquad \lambda \equiv \frac{\mu_i - \mu_j}{1 - \left(\alpha_i + \alpha_j\right)}$$

so that the gap between the two prices is a first-order auto-regressive process; the effect of a shock dies away geometrically; and it is possible to describe this decay by a single statistic. The time taken for half of the magnitude of a price difference to die away is referred to as the *half life*, defined as

$$(8) \qquad HL \equiv \frac{\ln\left(0.5\right)}{\ln\left(1 + \alpha_i - \alpha_j\right)} = \frac{\ln\left(0.5\right)}{\ln\left(1 + \boldsymbol{\gamma}\boldsymbol{\alpha}\right)}$$

---

[8] I.e. instead of using a vector auto-regression it is possible to create a single new variable (the price gap) and to estimate $\alpha_i + \alpha_j$ from an auto-regression (including a constant), If it is correct that the variables are cointegrated (with LOOP), then the variable $\left(p_t^i - p_t^j - \lambda\right)$ is stationary and it is not necessary to use the Dickey-Fuller distribution for hypothesis tests on $\widehat{\alpha_i + \alpha_j}$.

We discuss the optimal estimator of the half life in Appendix 5, where we show that it can adequately be estimated using the most obvious formula, namely $\ln(0.5)/\ln(1+\boldsymbol{\gamma}\hat{\boldsymbol{\alpha}})$.

The dispersion in prices depends partly on the magnitude of the shocks moving prices apart and partly on the speed with which the prices then return to equilibrium. The easiest way to see this decomposition is to consider a simplification of equation (6) where there are no lagged price changes (i.e. $\boldsymbol{\pi}^{(k)} = \mathbf{0}$). From equation (6) we derive:[9]

$$\mathsf{E}\left[\left(p_t^i - p_t^j\right)^2\right] = \mathsf{var}\left[p_t^i - p_t^j\right] + \left(\mu_i - \mu_j\right)^2$$

(9)

$$\underbrace{\frac{\exp\left\{-2\times \textsf{Half life}\right\}}{4}}_{\substack{\text{a function of the Half life}\\ \text{(or speed of adjustment)}}} \mathsf{E}\left[\left(p_{t-1}^i - p_{t-1}^j\right)^2\right] + \underbrace{\omega_{i.i} + \omega_{j.j}}_{\substack{\text{variance of}\\ \text{the disturbances}}} - \underbrace{2\omega_{i.j}}_{\substack{\text{covariance of}\\ \text{the disturbances}}} + \underbrace{\left(\mu_i - \mu_j\right)^2}_{\substack{\text{constant}\\ \text{terms}}}$$

The first line of this formula decomposes the variance of the price gap into a component due to variation in prices and a component dependent on the constant term $\left(\mu_i - \mu_j\right)^2 = \boldsymbol{\gamma\mu\mu}'\boldsymbol{\gamma}'$. The latter may represents the gap in prices which is permanent during the period of analysis, which may be due to problems with market integration or may represent constant quality differences. In Federico's terminology the permanent gap is an issue of LOOP, rather than market efficiency.

The second line of the formula is model-dependent and further decomposes $\mathsf{var}\left[p_t^i - p_t^j\right]$ into three components:

First, the variances of the disturbances $\omega_{i.i} + \omega_{j.j}$, which cause prices to change: larger disturbances result in greater dispersion of prices;

Secondly, the covariance of the disturbances $\omega_{i.j}$, which reduces the dispersion in prices. There are two ways of interpreting this variable: first, it is a measure of the correlation of the underlying shocks (e.g. if both towns suffer bad weather at the same time due to

[9] Since $HL = \ln(0.5)/\ln(1+\boldsymbol{\gamma\alpha})$ it follows that $1+\boldsymbol{\gamma\alpha} = 0.5\exp\{-HL\}$. In Appendix 6 we derive this result formally and discuss discuss the general case with lagged price changes.

weather being correlated); second, as we show in Appendix 7, it is the adjustment of prices which takes place within the period of observation (i.e. within-week price adjustment in our data). It is not possible to identify these two different effects.

Thirdly, the speed of adjustment (which Federico, 2012, refers to as efficiency): the shorter the half life, the less the price dispersion this period is due to prices being out of equilibrium in the previous period.

From our data set we are able to obtain estimates of these three components of price efficiency. The advantage of this is that we are able to ask, not just what correlates with reduced or increased price dispersion (Federico, 2011), but also the *mechanism* through which changes determine the price dispersion. For example, does improved transportation reduce the variance of domestic price shocks (by allowing imports to flood in to the domestic market more cheaply or quickly, and therefore offset changes in the domestic harvest)? Or does improved transportation increase the covariance of shocks (by linking regional markets more strongly to each other)? Or does improved transport increase the speed of adjustment of one market to another? Before we analyse these effects in more detail, however, we address some issues of estimation.

4. Cointegration estimation

4.1 Models estimated for the whole period

In this section we estimate the VECM models for the whole period 1770-1820. This means that our analysis is based on fifty years of weekly data and thus each regression contains approximately 2,500 observations, the precise number depending on the number of missing observations and the number of lagged dependent variables. The use of time series with so many observations potentially increases the efficiency of our estimates: conversely, there is considerable structural instability in the data and estimating models on data with such structural breaks may be mis-leading. We return to the issue of structural breaks in section 4.2.

To clarify our procedures, we illustrate many of our models using prices from Bedfordshire and Buckinghamshire for 1770-1820. Alphabetically, these counties are the first adjacent-county pair our data set (i.e. we chose them randomly, not because their prices series have any special characteristics). Geographically, they are large, adjacent counties; both of them are agricultural and both have almost complete data (just one missing observation each). For these two counties we create three different data sets: (i) the original weekly data; (ii) end-of-month data using the last price within each calendar month; (iii) annual data using the first price in October. Note that one problem with constructing monthly data is that our underlying data are "week-ending" data: sometimes there are five observations in a month and sometimes only four. So the monthly data are not equally spaced. We also estimated a model using every fourth observation so that the data were only approximately monthly but were exactly equally spaced: the results were almost identical to those from monthly data. Note also that the monthly and annual data are not within-period averages of the weekly data, so there is no issue of temporal aggregation as discussed in Taylor (2001).

From these three versions of the data we estimate the following models (regardless of the data frequency, the half lives are measured in weeks):

$$\text{annual data} \quad \widehat{\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix}} = \begin{bmatrix} -0.603 \\ 0.272 \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} -0.010 \\ 0.015 \end{bmatrix} \quad \begin{matrix} HL = 17.4 \text{ weeks} \\ EHL = 19.0 \text{ weeks} \end{matrix}$$

$$\text{(10)} \quad \text{monthly data} \quad \widehat{\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix}} = \begin{bmatrix} -0.256 \\ 0.249 \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} -0.006 \\ 0.009 \end{bmatrix} \quad \begin{matrix} HL = 4.3 \text{ weeks} \\ EHL = 4.3 \text{ weeks} \end{matrix}$$

$$\text{weekly data} \quad \widehat{\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix}} = \begin{bmatrix} -0.127 \\ 0.133 \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} -0.004 \\ 0.004 \end{bmatrix} \quad \begin{matrix} HL = 2.3 \text{ weeks} \\ EHL = 2.3 \text{ weeks} \end{matrix}$$

Including seasonal dummies for the monthly and weekly data makes no quantitative difference. It is notable that using higher frequency data results in much shorter estimated half lives: using weekly data, rather than monthly, results in a half life of two weeks, rather than four.

Figure 5 about here (distribution of half lives)

The results from Bedfordshire and Buckinghamshire are fairly representative of other pairs of counties. We estimate the half life for each county pair and illustrate our results in Figure 6 (further description is in Appendix 8). On average, the half life estimated from data measured at an annual frequency is twenty weeks, skewed heavily to the right; whereas the half life estimated using weekly data is about eight weeks, with much less skew. The second panel of Figure 6 uses half lives estimated from weekly data to compare the distribution of half lives for all counties and for adjacent counties. Markets appear more efficient (have a shorter half life) when counties are adjacent: the average half life is only four weeks, instead of eight.

From this we conclude that it is possible to get dramatically different estimates of the half life by using data of different frequencies. Although there is parameter instability in all of the models, this is unlikely to explain the differences in half life estimates entirely, since all three regressions are based on the same span of data (i.e. 1770-1820) and suffer the same instability. The more important problem is that the weekly models display significant serial correlation, suggesting that the VAR of equations (4) and (5) does not fit the data if we impose the restriction $\boldsymbol{\pi}^{(k)} = \mathbf{0}$. All of the models in (9) have this restriction; all have biased parameter estimates and the degree of bias depends upon the frequency of the data used in estimation.

This analysis underlines the fact that one cannot compare half lives from models estimated on data of different frequencies. So research based on annual data (Sharp and Weisdorf, 2013; Studer, 2008), is not comparable to research using monthly data (Bateman, 2011; Buyst, Dercon and Van Campenhout, 2006; Goodwin and Grennes, 1998; Goodwin, Grennes and Craig, 2002; Jacks, 2005; Marks, 2010; Trenkler and Wolf, 2005), which is not comparable to research using data with two observations per month (Ejrnæs and Persson, 2000), which is not comparable to research using weekly data (Ejrnæs and Persson, 2010; Federico, 2007; Hynes, Jacks and O'Rourke, 2012). One possible solution to this would be for authors to report half lives based on both their underlying data and also from the same data sampled at a lower frequency, although the latter would only be an imperfect measure for comparison purposes.

Interestingly, estimated half lives from annual data appear slightly longer for adjacent counties than for all counties, suggesting that attempts to correlate market efficiency with distance may be ineffective or misleading when data are measured at low frequencies. If this result could be generalised then it might explain why Studer (2008, Table 5) finds only weak or ambiguous correlation between market efficiency and distance.[10]

There is no reason to believe that the simple VECM model (i.e. with no lagged price changes) is a suitable model to explain prices and this is confirmed in our data both by

---

[10] For example, Studer's average estimate of $\alpha_i - \alpha_j$ for 1870-1914 is −0.46 when the distance is 150-300 km (a half life of 1.12 years); when the distance is 600-1000 km, the adjustment is −0.60 (a half life of 0.76 years), which is considerably faster. But Studer is using annual averages (p. 396), so the half lives are all biased up (Taylor, 2001).

the presence of serial correlation in the disturbances and the fact that lagged price changes are statistically significant when included.

Several authors include additional lags in the VAR (Persson, 1999, ch.5; Bateman, 2011, whose procedure is explained in Bateman, 2007; Marks, 2010); but they do not plot the full impulse response function and appear to measure market efficiency using the loadings alone, despite the fact that this gives no meaningful description of the response of prices to market disequilibrium. Trenkler and Wolf (2005) estimate a VAR with more lags, but then re-estimate the model with just one lag to get a half life. Goodwin and Grennes (1998) and Goodwin, Grennes and Craig (2002) and Ejrnæs, Persson and Rich (2006) illustrate the effects of shocks on different markets on the full set of prices but do not provide a measure of the speed of convergence.

To illustrate the effect of including lagged price changes, we return to the Bedfordshire and Buckinghamshire prices, using weekly data. Two sample models that we estimated are (to save space we do not report the estimated constant and seasonal dummies):

$$\widehat{\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix}} = \begin{bmatrix} -0.095 \\ 0.099 \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} -0.005 & 0.257 \\ 0.211 & -0.034 \end{bmatrix} \begin{bmatrix} \Delta p_{t-1}^i \\ \Delta p_{t-1}^j \end{bmatrix} \tag{11}$$

$$\widehat{\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix}} = \begin{bmatrix} -0.082 \\ 0.081 \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} -0.039 & 0.289 \\ 0.242 & -0.077 \end{bmatrix} \begin{bmatrix} \Delta p_{t-1}^i \\ \Delta p_{t-1}^j \end{bmatrix} + \begin{bmatrix} -0.071 & 0.078 \\ 0.112 & -0.062 \end{bmatrix} \begin{bmatrix} \Delta p_{t-2}^i \\ \Delta p_{t-2}^j \end{bmatrix} \tag{12}$$

Comparing (11) and (12) with the weekly-data version of equation (10), it is apparent that the loadings get smaller as more lags are included. If an attempt were made to estimate the half life just from the loadings from equations (11) and (12), regressions with more lags would suggest longer half lives, illustrated in the first row of Table 1. This table also contains results for more lags, going up to 53 weeks, to take account of any additional seasonal effects. The disadvantage of including so many lags is that the confidence intervals (not reported here) are much wider.[11]

[11] We do not address the issue of optimal lag length in this paper. Conventional criteria, such as information criteria, typically choose a compromise to maximise goodness of fit subject to minimising the number of explanatory variables. Which criterion is optimal is sensitive to the objectives of the research (so estimation, testing and forecasting

Table 1 about here

However, in this instance the loadings now under-state market efficiency. Consider a hypothetical situation where, from market equilibrium, prices diverge due to a disturbance causing a rise in Buckinghamshire prices in period 1 while Bedfordshire prices are constant. From equation (11) the prices move towards each other in period 2, not just due to the error correction term, but also because – in that period – the Bedfordshire price rises by $0.257\Delta p_1^{Buck}$ (there is also a related change in the Buckinghamshire price). In addition to the decrease in the disequilibrium of 0.194 from the error correction, there is an additional 0.291 from the effect of the lagged price changes. These effects are illustrated in the second row of Table 2.

Figure 6 about here (impulse response functions Beds-Bucks)

In Figure 6 we plot the impulse response functions for Bedfordshire and Buckinghamshire from models with differing lags of price changes: since the data are weekly, we consider up to 53 lags to allow for seasonal effects (although the model also includes seasonal dummy variables, which make little difference). We provide details of the construction of these impulse response functions in Appendix 5.2. The solid black line shows the model estimated with no lags and demonstrates geometric decay, i.e the disequilibrium decays at the same speed regardless of the length of time since the disturbance. So long as two or more lags are included the impulse response function is more-or-less the same: the shape is quasi-hyperbolic, with relatively fast decay for the first few weeks and thereafter relatively slow decay.

When there are lagged price changes and the decay in the price gap is not geometric, there is no single measure that can summarise the speed of adjustment: any measure that we use will only crudely measure the adjustment pricess. The method we choose is to continue to use the half life which we do so by simple linear interpolation (where the graph cuts the horizontal line). When more than two lags are included in the estimation, the half life ranges from 1.23 to 1.04.

The third and fourth rows of Table 2 show that the example cannot be generalised: if we look at all 780 county pairs then the mean average half life is actually *higher* when

---

might all yield different answers); for the purpose of this paper more lags will generally be better than fewer.

several lags are included (the same is true for the median). However, it is the case for adjacent-county pairs that including more lags results in shorter average half lives. Figure 8 shows the distributions of the half lives for all 780 county pairs for differing lags. Adding a few lags results in longer half lives and adding very large numbers of lags results in slightly shorter half lives. So, although the inclusion of additional lags changes the conclusions, the nature of the change is ambiguous.

4.2 Models estimated on sub-samples of the data

Our models hitherto have all been estimated on the whole sample from 1770 to 1820. This is obviously inappropriate if there is parameter instability, especially since the economic issue is potential changes in efficiency. One way to approach this problem would be to look for structural breaks in each time series. But a problem with this is that small breaks might not appear statistically significant; also, we are testing for breaks in a variety of different parameters (including the variance and covariance of the disturbances). Our preferred solution is to divide the data set into 4650 sets of weekly data for a given harvest year for each adjacent county pair: so, for example, one data set would be the relevant weekly observations for Bedfordshire and Buckinghamshire for the harvest year 1780-81. This method is analogous to that of Jacks (2011), who looks at bilateral price comparisons within the year.

We now start by looking at the first line of the decomposition in equation (9), with one difference: because variances are difficult to interpret we look at the absolute difference between the mean prices rather than $\left(\mu_i - \mu_j\right)^2$ and the standard deviation of the price gap rather than the variance:

$$\text{(13)} \qquad \begin{aligned} \text{abs.diff}\left[p^i, p^j\right]_y &= 100 \times \left|\overline{p}^i_y - \overline{p}^j_y\right| \qquad \overline{p}^i_y \equiv n^{-1}\sum_{w\in y} p^i_w \\ \text{st.dev.}\left[p^i - p^j\right]_y &= 100 \times \sqrt{n^{-1}\sum_{w\in y}\left(\left(p^i_w - \overline{p}^i_y\right) - \left(p^j_w - \overline{p}^j_y\right)\right)^2} \end{aligned}$$

In both cases we have multiplied by one hundred so that the figures are percentages. To illustrate the resulting 4,650 statistics (93 adjacent-county pairs and fifty years) that we have calculated, we plot the mean average of both statistics for each harvest year in Figure 7.

Figure 7 about here (abs diff and st dev)

Except for the spike in both series after the Napoleonic Wars, there is no clear trend downwards, consistent with the summary statistics presented in section 2.

We then estimate the model of equation (6) in each of the 4650 within-year data sets and thus obtain a panel of estimated parameters. Although these estimates are unlikely to be highly efficient estimates of the true parameters, we have sufficiently many that our further panel analysis will still be efficient.

One problem with this approach is that there is a strong seasonal pattern in prices that is variable over time, and we are unable to model seasonal effects when using data within a single year. However, Brunt and Cannon (2002) show that the seasonal pattern is approximately saw tooth: in about the 33rd week of the year, at harvest time, prices fall dramatically until about the 45th week: thereafter they rise approximately exponentially (so log prices rise linearly). From this stylised fact we use the forty observations from the 45th week of year y to the 33rd week of year $y+1$ and ignore seasonal effects (the stochastic trend is modelled through the constant term, which is not restricted to lie in the cointegrating space). We refer to the parameter estimates for this year as belonging to year y. Forty data points is a relatively small number of observations, and we lose observations due to the need for lagged variables and due to missing data: where there are fewer than thirty observations we do not estimate parameters at all.

We measure the magnitude of the shocks using their average estimated standard deviation. So for county-pair *i, j* we use

$$\text{magnitude of shocks}\left(i,j;y\right) \equiv 100 \times \frac{\sqrt{\hat{\omega}_{i.i,\text{year } y}} + \sqrt{\hat{\omega}_{j.j,\text{year } y}}}{2} \qquad (14)$$

where we multiply by one hundred so that the figures can be interpreted as percentages (note that the estimator of $\hat{\omega}_{i.i}$ depends on how many lags are included in model 6). We illustrate the annual average magnitudes for this measure in Figure 8. The vertical axis is measured in percentages, so for the first part of the period this measure of price dispersion was about 4 per cent. From about 1793 onwards it rose, coinciding with the Revolutionary and Napoleonic Wars. This is consistent with Jacks (2011), but not consistent with Figure 2, which showed that the standard deviation of all prices did not rise. The apparent contradiction is resolved by the observation that relatively close markets became less integrated while overall dispersion of prices did not rise. This suggests that any effect of higher volatility from the Revolutionary and Napoleonic Wars was masked, or even offset, by other factors.

To measure the correlation of the shocks we use

$$(15) \qquad \text{correlation of shocks}\left(i, j; y\right) \equiv \frac{\hat{\omega}_{i.j,\text{year } y}}{\sqrt{\hat{\omega}_{i.i,\text{year } y} \times \hat{\omega}_{j.j,\text{year } y}}},$$

which is illustrated in Figure 9.

Figures 8 and 9 about here (st dev and correlation)

One problem with these two summary measures is that they do not show when one shock has a larger magnitude than the other. Although it does not fit neatly within our decomposition in equation (9), we also consider a measure of the relative size of the two shocks, namely the larger standard deviation divided by the smaller:

$$(16) \qquad \text{relative size of shocks}\left(i, j; y\right) \equiv \frac{\max\left[\sqrt{\hat{\omega}_{i.i,\text{year } y}}, \sqrt{\hat{\omega}_{j.j,\text{year } y}}\right]}{\min\left[\sqrt{\hat{\omega}_{i.i,\text{year } y}}, \sqrt{\hat{\omega}_{j.j,\text{year } y}}\right]}.$$

This measure is illustrated in Figure 10. Although there is a very slight downward trend in the series (suggesting shocks were becoming more similar in size), this is dwarfed by the idiosyncratic changes from year to year.

Finally we use the half life (defined in equation 8 and discussed in Appendix 5) as a measure of the speed of adjustment.

Figures 10 and 11 about here (ratio and half life)

Figure 8 confirms that the one of the major causes of the greater price dispersion illustrated in Figure 7 was that the disturbances were larger: the peaks in price disturbance in Figure 8 coincide with peaks in Figure 7, although the magnitudes are not necessarily the same. This is prima facie evidence that prices became more dispersed, not due to declining efficiency of the market, but due to the shocks hitting the economy. However, Figure 9 shows that over time the disturbances to markets became more correlated and this attenuates the effects of larger shocks on price dispersion. Higher correlation does not mean that the disturbances became more similar in size, and so we look also at the ratio of the magnitude of the shocks in Figure 9: this is the ratio of the standard deviation of the more variable disturbance to that of the less variable disturbance. If the Revolutionary and Napoleonic Wars resulted in more similar shocks (i.e. a source of additional shocks that was the same for all markets) then the variances of the shocks should have become more similar. Over the whole period, when the magnitude of shocks is high, they are both more correlated (correlation of 0.73) and the relative size falls (correlation of -0.46).

It is notable that figures 8, 9 and 10 show results that are almost identical regardless of the number of lags in the VECMs, suggesting that these measures are relatively robust to the precise model used. The half life, however, is sensitive to the estimation method used. In all cases, however, the half lives tend to fall over the period: regressing the average half life on a trend results in a coefficient of about -1 per cent; this is true even if the estimation is only for the period 1792-1815. This is only an informal calculation: but, using Newey-West standard errors to compensate for the obvious serial correlation, it suggests that the relationship is statistically significant at the 5% level when the half lives are calculated from VECMs with zero, two or three lags and at the 10% level for one lag. This suggests that market efficiency was improving throughout the period including the Revolutionary and Napoleonic Wars. The reason that this does not show up in measures of price dispersion is that the shocks to the economy were simultaneously increasing.

To summarise this section: we have shown informally how the dispersion in prices can be decomposed into the magnitude of the shocks, the correlation of the shocks and the speed to convergence (half life). Estimates of market efficiency can be highly sensitive to both the frequency of the data and the number of lags included in time series models. However, the component of market efficiency that is most sensitive is the measure of between-period arbitrage (i.e. the half life, derived from the impulse response function) while estimates of the variance and covariance of the shocks are much less affected. The differences are sufficiently large that they suggest that comparison of research using different methods or frequencies is hazardous.

Our data confirm that the Revolutionary and Napoleonic Wars saw increased price dispersion, but we show that this was not due to less efficient markets. The evidence suggests that market efficiency continued to increase, even while the magnitude of shocks grew larger: the reason for greater price dispersion was that the latter predominated.

5. The effect of transport on market efficiency

In the previous section we showed that there was evidence that market efficiency improved, but that this failed to reduce dispersion in prices because the magnitude of the shocks hitting the economy were simultaneously increasing. This raises the question of whether we can find any effect of transport and communication variables on market efficiency.

Our procedure is similar to that of Jacks (2011), but for three differences. First, we consider only adjacent county pairs. This is mainly because transport – such as roads or canals – is only conceptually easy to measure for adjacent counties: where counties are not adjacent it is not obvious how they would be linked for arbitrage purposes. One linkage possibility is coastal traffic; but, to the extent that this is constant, it is already modelled in the fixed effect. We are also concerned about the statistical properties of using all 780 county pairs: since these are based on only 40 price series, they are not independent.

Second, Jacks looks at a single measure of price dispersion – albeit a slightly different one to us – whereas we look at the components of price dispersion.

Third, we increase the number of controls by the using both year and county-pair fixed effects, instead of the alternative of time-series variables that Jacks uses (such as severity of war, measured by battle casualties). The reason for this choice is that we are primarily interested in the effect of transport variables (we take it as read that warfare disrupted markets) and so are content to use a relatively large set of control variables.

We use two transport variables. The first is a dummy variable indicating that the two counties were linked by a canal. The second is a measure of turnpiked roads in the two counties defined as

$$\text{Road}_{(i,j).t} \equiv \frac{M_{i.t} + M_{j.t}}{A_i + A_j} \tag{17}$$

where $M_{i.t}$ is the mileage of turnpiked road in county *i* in year *t* and A_i is the area of the county in hundreds of square miles. At first it might seem strange to measure road linkages by the average road density, since transport links are typically thought of as *between* two markets. But recall that our data are average prices within counties and therefore for we would only expect one county's average price to converge to the other's if all markets were connected within the two counties. Given the price and road data that are available, this measure seems appropriate.

A final consideration is that market integration might have improved due to a reason other than improved transport. Since grain holders could arbitrage across time as well as across space, the arrival of news might have been equally or more important than the speed or cost of transport. We attempt to measure this by using a measure of communication. We use data on newspaper circulation in the towns from which our wheat prices were collected. Underlying data on newspaper circulation were taken from Gibson (1991) and from this we calculated the proportion of towns in adjacent-county pairs that had at least one newspaper. As a robustness check we also considered the average number of newspapers in circulation: the results were quantitatively very similar.

Table 2 about here (regression results lag zero)

Our first set of regression results are reported in Table 2. The first column reports the regression for the permanent price gap (measured in equation 13.a) which is a measure of violation of LOOP. From Figure 7 we know that there was no trend in this variable: it also appears that Canals, Roads and Newspapers had little effect on it. One possible reason for this is that the permanent price gap reflects unchanging regional quality differences and therefore observed prices would not be equal even with perfect arbitrage.

The second column reports the regression for the standard deviation of the price gap, which we know to have increased during the Napoleonic Wars. Given the huge variability in price dispersion – and the fact that road and canal and newspaper networks evolved only relatively slowly – it would be unsurprising if none of the variables were statistically significant. However, the Canal indicator is statistically significant at conventional levels and it suggests that a canal reduced the root-mean squared price difference by one-quarter of one per cent. The effect for newspapers is statistically significant at the ten per cent level, but the effect is relatively small: in a county pair with a total of ten towns, the presence of one additional newspaper in a town previously without a newspaper would reduce the root-mean-squared price difference by only 0.7 per cent.

The remaining four columns of Table 2 report the regression results for the components of market integration based upon estimates of the VECM with no lags. From section 4.2 we know that the first three measures were very similar regardless of the number of lags included in the VECM and this conclusion continues here: analogous regressions with the statistics calculated using more lags are similar (we report the results of those regressions in Appendix 8). Both Roads and Canals appear to reduce the magnitude of the shocks:[12] an extra ten miles of turnpike per hundred square miles would reduce the standard deviation of the shocks by about one-third of one per cent, while the presence of a canal would reduce the standard deviation by about one-sixth of a per cent. Theoretically, the effect of transport on the variance of price changes is ambiguous (depending on elasticities of supply and demand); but it appears in this instance that the lower transport costs allowed risk-sharing through pooling of risks in separate locations. The effect is large, as evidenced by the R-squared.

Our other two measures of the shocks are the ratio of the magnitudes and the correlation. Roads and Newspapers appear to reduce the ratio of the variance of the shocks: in other words, if a shock hits one market then the size of the shock hitting the other market is more likely to be the same size. This is *prima facie* evidence that both Roads and Newspapers increase market efficiency, as the disturbances in the two markets have a more similar magnitude. Surprisingly the Roads variable has only a minimal effect on the correlation of the disturbances, but Newspapers have a large and statistically significant effect, suggesting that they explain within-week price adjustment (Appendix 7). The Canals variable is only marginally statistically significant..

Since our estimates of the half life are sensitive to the number of lags in the VECM, we report results for different lag lengths in Table 3 to facilitate comparison. Were we to look at estimates of the half life based on a VECM with no lagged price changes we would conclude that both Roads and Canals had a positive and statistically significant effect on the half life, which suggests that they reduce market efficiency. The effects appear to be large: an extra ten miles of road per hundred square miles apparently increases the half life by two-thirds of a week (i.e. four to five days) and a canal by one-third (two days). But analysis of half-life estimates based on different lag lengths would result in quite different conclusions as the effects of all three explanatory variables are very imprecisely estimated when the VECM includes even one lagged price change. On this evidence the variability in our half life estimates is too large to be able to draw any meaningful inferences about the effects of transport or information on market adjustment.

[12] The results for the Road variable are slightly sensitive to the number of lags in the first-stage VECM: when there is one lag the t-statistic on the Road variable falls to 1.94 with a p-value of 0.052.

Table 3 about here (comparative regression results)

In terms of our understanding of market efficiency in the Revolutionary and Napoleonic Wars, we conclude that – although prices converged to equilibrium more quickly – we are unable to explain why. Our transport and communication variables seem to have had more effect on price changes at a frequency of less than a week, thus raising the measured correlation of county shocks.

6. Summary and Discussion

We have analysed the comprehensive data set of London Gazette English grain prices for 1771-1820. In the spirit of Federico (2012), who notes that different authors have used different techniques, we have reported a variety of measures. Summary statistics of the data set, such as the coefficient of variation suggested by Federico (2011), suggest that the market was remarkably stable over this fifty year period, despite the expansion of transport networks and the shocks of wars. Looking at the graphs in section 2, it is difficult to see anything that has changed over the period, other than prices all moving up during the Napoleonic wars. In this paper we have attempted to see whether this is due to a genuine absence of change or whether different changes approximately cancel each other out.

An increasingly popular tool for measuring market efficiency is the use of VECM models. Whilst it is well understood that the conclusions of these models depend upon the data frequency, we have – until now – had little idea of the magnitude of this effect. Since we have a complete set of weekly data, we have been able to estimate the speed of convergence to equilibrium, not only on high frequency data but also on lower frequency data, and thus quantify the importance of this issue. In Section 4.1 we show that weekly data with a half life of about eight weeks would appear to have a half life of eleven weeks if the data were sampled at a monthly frequency and twenty weeks if the data were sampled annually. Previous studies of market integration suggest that prices adjusted very slowly: our results here suggest that estimates of the speed of adjustment may have been too pessimistic.

A further issue that we examine is whether the underlying assumption of geometric decay in price dispersion is correct. Using models with richer dynamic structures, we find that the convergence to equilibrium is quasi-hyperbolic, rather than geometric; that estimates of the half life may differ significantly; and that this may change the ordering of which markets we believe to be most efficient. This may be because price behaviour was better modelled with a TECM rather than an ECM approach but, since we have only averages of prices from different markets, it would not make sense to implement a TECM with our data.

These two points taken together suggest that it may be difficult to compare reliably previous studies that use different frequency data or omit additional lags in time series estimation.

Our analysis supports the work of Jacks (2011) and Dobado-González, García-Hiernaux & Guerrero (2012), who find that prices became more dispersed during the Napoleonic Wars (although price dispersion also remained high immediately after the conflict was over in 1816-17). This was not due to the breakdown of the Law of One Price (LOOP): the permanent price gaps between counties show no secular trend. The major reason for the increase in price dispersion was disturbances in the price dynamics: shocks from abroad mattered more, and so the disturbances to prices became larger, more highly correlated and more similar in size. There was an increase in market efficiency, as measured by the half life, but the effect of this was relatively small.

Our final contribution is to see whether transport and communication variables can explain either the overall behaviour of prices or the underlying components. The transport variables, but not our measure of newspapers, reduce the magnitude of random changes in prices, suggesting that arbitrage acts as a form of risk-pooling and reduces overall price variation. So the primary importance of the transport variables appears to have been on the magnitude of the shocks, although this was not the only mechanism. Newspapers sped up the transmission of information, so that shocks to prices were more correlated: information arrived in different places at the same time (at least, within the same week).

Market efficiency (moving towards equilibrium) occurs both within the period of observation (i.e. within the week) and over longer periods: the latter is measured through the half life. We have some evidence that the half life fell over the period 1770-1820, but estimates of this variable are sensitive to the model used: regardless of this, we are unable to explain the decline in the half life with the transport and communication variables that we have used here.

References

Albert, W. (1972) *The turnpike road system in England, 1663-1840* (Cambridge: University Press).

Aldcroft, D.H., and Freeman, M.J. (1983) *Transport in the Industrial Revolution* (Manchester: Manchester University Press).

Bagwell, P.S. (1974) *The transport revolution from 1770* (London: Batsford).

Bateman, V.N. (2007) "The evolution of markets in early modern Europe, 1350-1800: a case study of wheat prices" University of Oxford Department of Economics Discussion Paper No. 350.

Bateman, V.N. (2011) "The evolution of markets in early modern Europe, 1350-1800: a case study of wheat prices" *Economic History Review 64(2)*, pp. 447-471.

Bogart, D. (2005a) "Did turnpike trusts increase transportation investment in eighteenth-century England?" *Journal of Economic History, 65(2)*, pp. 439-468.

Bogart, D. (2005b) "Turnpike trusts and the transportation revolution in 18th century England" *Explorations in Economic History 42,* pp. 479–508.

Bogart, D. (2009) "Turnpike trusts and property income: new evidence on the effects of transport improvements and legislation in eighteenth-century England" *Economic History Review, 62(1),* pp. 128-152.

Bogart, D. (2011) "Did the Glorious Revolution contribute to the transport revolution? Evidence from investment in roads and rivers" *Economic History Review, 64(4),* pp. 1073-1112.

Buyst, E., Dercon, S., and Van Campenhout, B. (2006) "Road expansion and market integration in the Austrian Low Countries during the second half of the 18th century" *Histoire & Mesure, 21(1),* pp. 185-219.

Brunt, L. and Cannon E.S. (2002) "Do banks improve financial market integration?" University of Bristol, mimeo.

Brunt, L. and Cannon, E.S. (2013) "The truth, the whole truth, and nothing but the truth: The English Corn Returns as a data source in economic history, 1770-1914" forthcoming *European Review of Economic History*..

Chartres, J.A., and Turnbull, G.L. (1983) "Road transport"; chapter 3 in Aldcroft and Freeman (1983).

Dobado-González, R., García-Hiernaux, A., & Guerrero, D.E. (2012) "The integration of grain markets in the eighteenth century: early rise of globalization in the west" *Journal of Economic History, 72(3)*, pp. 671-707.

Donaldson, D. (2010) "Railroads of the Raj: estimating the impact of transportation infrastructure" NBER Working Paper 16487.

Donaldson, D., and Hornbeck, R. (2012) "Railroads and American economic growth: a "market access" approach" MIT Working Paper.

Duckham, F. (1983) "Canals and river navigations"; chapter 4 in Aldcroft and Freeman (1983).

Ejrnæs, M. (1999) "Appendix to chapter 5" in Persson, K.G. (1999) *Grain markets in Europe, 1500-1900* (Cambridge: Cambridge University Press).

Ejrnæs, M., and Persson, K.G. (2000) "Market integration and transport costs in France 1825–1903: a threshold error correction approach to the law of one price" *Explorations in Economic History, 37*, pp. 149–173.

Ejrnæs, M., and Persson, K.G. (2010) "The gains from improved market efficiency: trade before and after the transatlantic telegraph" *European Review of Economic History, 14(3),* pp. 361–381.

Ejrnæs, M., Persson, K.G., and Rich, S. (2008) "Feeding the British: convergence and market efficiency in the nineteenth-century grain trade" *Economic History Review, 61(S1)*, pp. 140–171.

Federico, G. (2007) "Market integration and market efficiency: the case of 19th century Italy" *Explorations in Economic History, 44,* pp. 293-316.

Federico, G. (2011) "When did European markets integrate?" *European Review of Economic History, 15*, pp. 93-126.

Federico, G. (2012) "How much do we know about market integration in Europe?" *Economic History Review, 65(2),* pp. 470–497.

Fogel, R. W. (1964) *Railroads and American economic growth: essays in economic history* (Baltimore: Johns Hopkins University).

Gerhold, D. (1993) "Packhorses and wheeled vehicles in England, 1550–1800" *Journal of Transport History, 14,* pp. 1-26.

Gerhold, D. (1996) "Productivity change in road transport before and after turnpiking, 1690-1840" *Economic History Review, 49(3)*, pp. 491-515.

Gibson, J.S.W. (1991) *Local newspapers, 1750-1920: England and Wales, the Channel Islands and the Isle of Man: a select location list* (Birmingham: Federation of Family History Societies).

Goodwin, B.K., and Grennes, T.J. (1998) "Tsarist Russia and the world wheat market" *Explorations in Economic History, 35,* pp. 405-430.

Goodwin, B.K., Grennes, T.J., and Craig, L.A. (2002) "Mechanical refrigeration and the integration of perishable commodity markets" *Explorations in Economic History, 39*, pp. 154–182.

Hynes, W., Jacks, D.S., and O'Rourke, K.H. (2012) "Commodity market disintegration in the interwar period" *European Review of Economic History, 16,* pp. 119–143.

Jacks, D.G. (2011) "Foreign wars, domestic markets: England, 1793–1815" *European Review of Economic History, 15*, pp. 277–311.

Leunig, T.C. (2010) "Social savings" *Journal of Economic Surveys,24(5)*, pp. 775–800.

Marks, D. (2010) "Unity or diversity? On the integration and efficiency of rice markets in Indonesia, c. 1920–2006" *Explorations in Economic History, 47,* pp. 310-324.

Maw, P. (2009) "Water transport in the industrial age: commodities and carriers on the Rochdale canal, 1804-55" *Journal of Transport History, 30(2),* pp. 200-228.

Moran, P.A.P. (1950) "Notes on continuous stochastic phenomena" *Biometrika 37(1),* pp. 17–23.

Ó Gráda, C., and Chevet, J-M. (2002) "Famine and market in *Ancien Régime* France" *Journal of Economic History, 62(3),* pp. 706-733.

Pawson, E. (1977) *Transport and economy: the turnpike roads of eighteenth century Britain* (London: Academic Press).

Persson, K.G. (1999) *Grain markets in Europe, 1500-1900* (Cambridge: Cambridge University Press).

Petersen, C., edited by Jenkins, A. (1995) *Bread and the British economy c. 1770-1870* (Aldershot: Scholar Press).

Phillips, P.C.B. (1991) "Error correction and long-run equilibrium in continuous time" *Econometrica, 59(4)*, pp. 967-980..

Priestley, J. (1831) *Historical account of the navigable rivers, canals, and railways of Great Britain* (London: Longman, Rees, Orme, Brown & Green).

Ravallion, M. (1986) "Testing market integration" *American Journal of Agricultural Economics, 68*, pp. 102-9..

Sharp, P., and Weisdorf, J. (2013) "Globalization revisited: market integration and the wheat trade between North America and Britain from the eighteenth century" *Explorations in Economic History, 50,* pp. 88-98.

Studer, R. (2008) "India and the great divergence: assessing the efficiency of grain markets in eighteenth- and nineteenth-century India" *Journal of Economic History, 68(2),* pp. 393-437.

Taylor, A.M. (2001) "Potential pitfalls for the purchasing-power-parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price" *Econometrica, 69 (2)*, pp. 473-498.

Timmins, G. (2005) "Paving the way: advances in road-building techniques in Lancashire, 1770-1870" *Journal of Transport History, 26(1),* pp.19-40.

Trenkler, C., and Wolf, N. (2005) "Economic integration across borders: the Polish interwar economy 1921–1937" *European Review of Economic History, 9(2)*, pp. 199-231.

Van Bochove, C. (2008) *The economic consequences of the Dutch* (Amsterdam: Aksant Academic Publishers).

Ward, J.R. (1974) *The finance of canal building in eighteenth century England* (Oxford: Oxford University Press).

Figures and Tables

Figure 1: Wheat prices 1770-1820

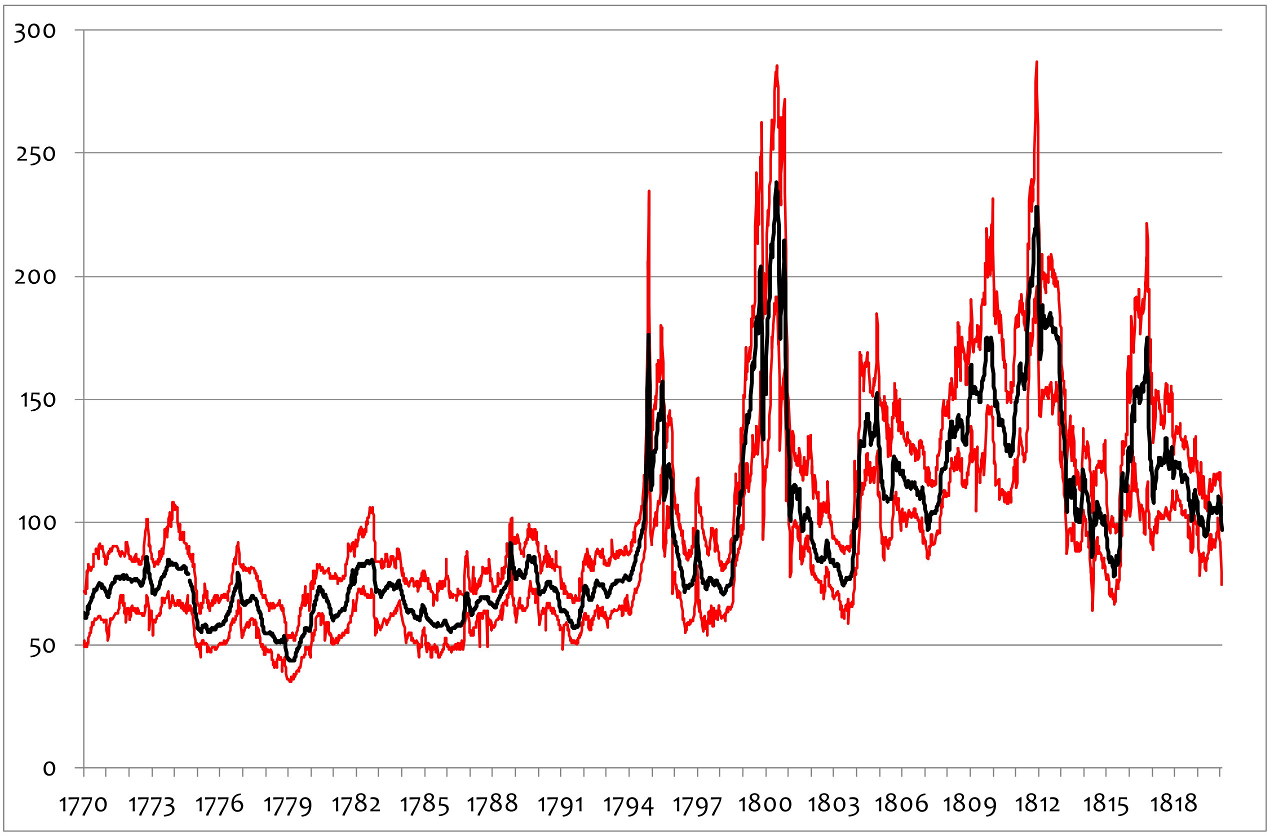
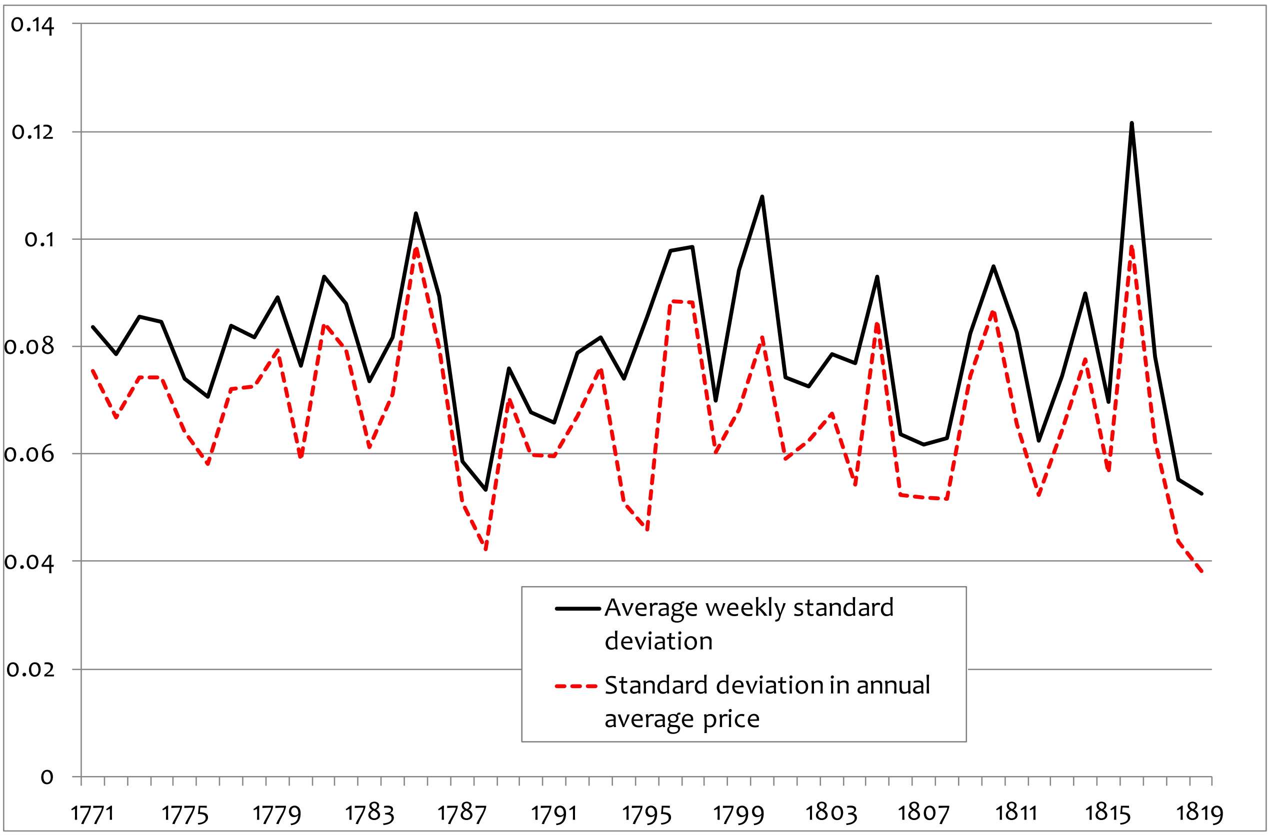


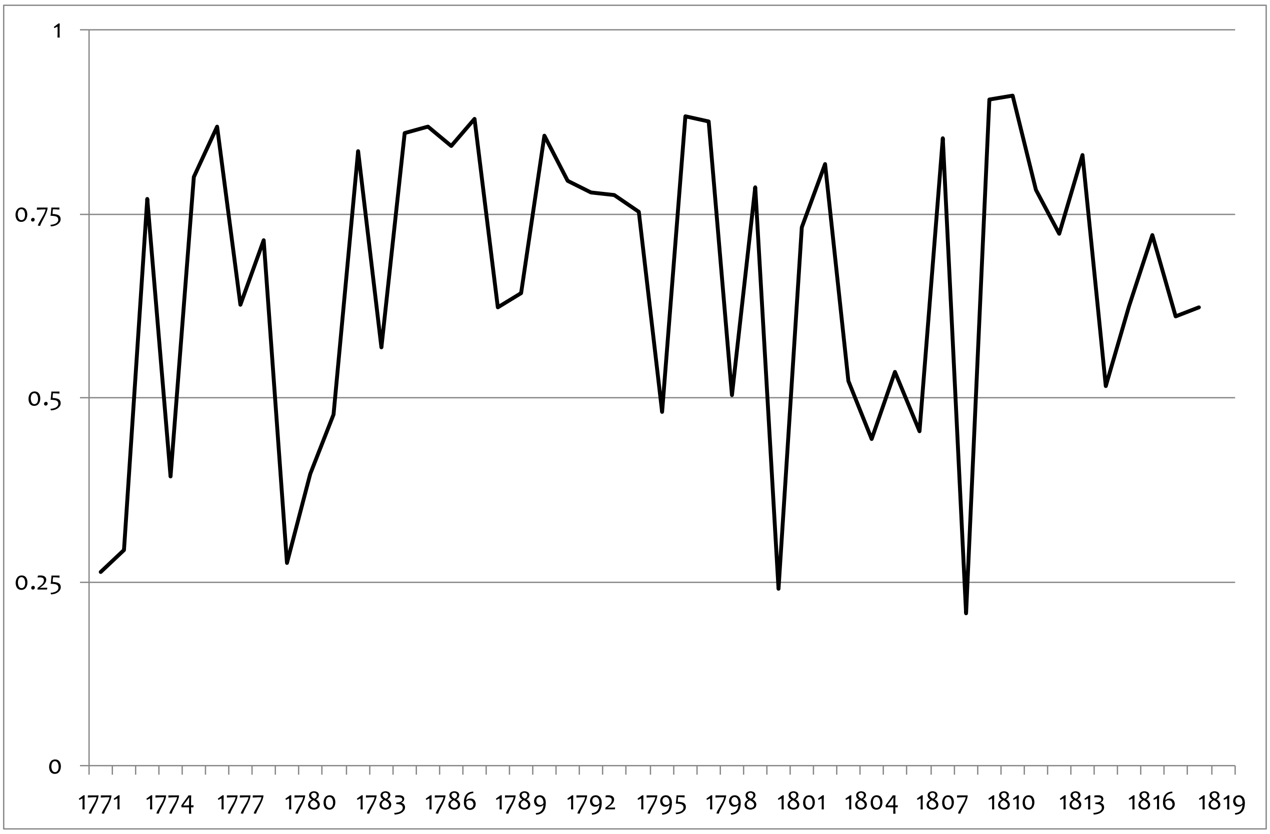
Figure shows the minimum, maximum and average London Gazette wheat price in each week from November 1770 to September 1820.

Figure 2: Dispersion of prices between counties



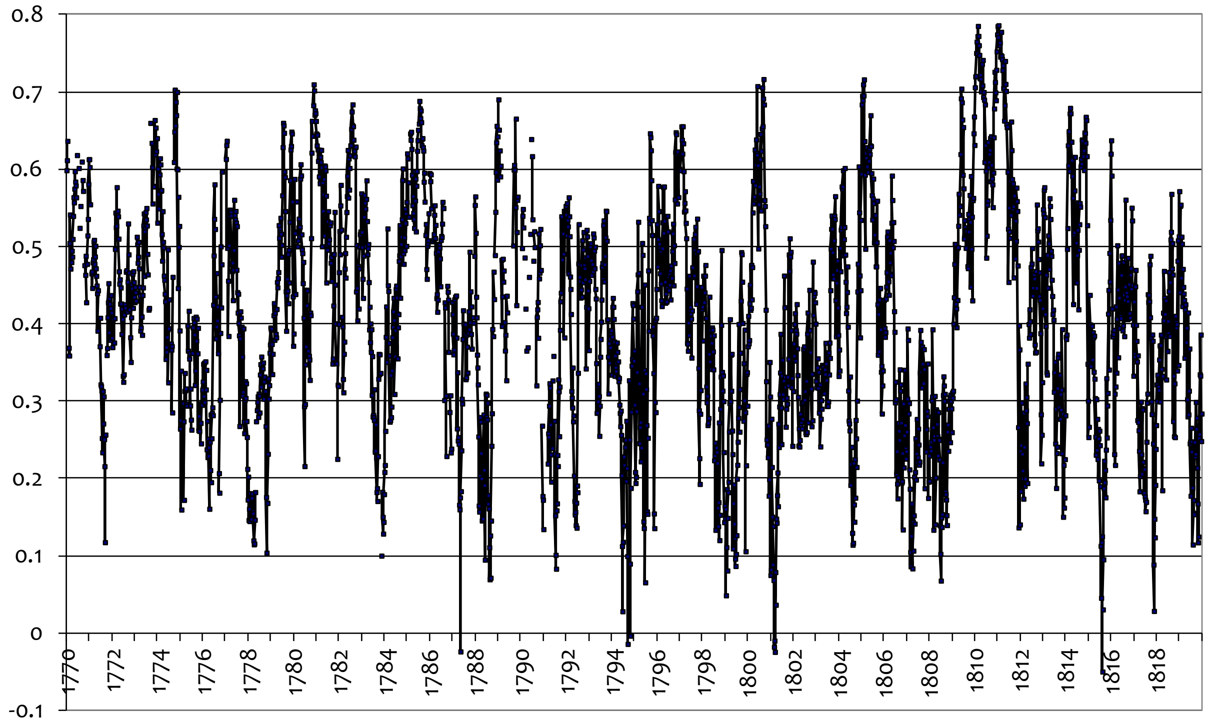
Average weekly standard deviation: the standard deviation of log prices is calculated for each week of the sample and then the 52 standard deviations are averaged for a harvest year (October-September). Standard deviation in annual average: the harvest-year mean price is calculated for each county and then the standard deviation is calculated of the forty mean prices.

Figure 3: Year-on-year correlations of cross-sections of prices



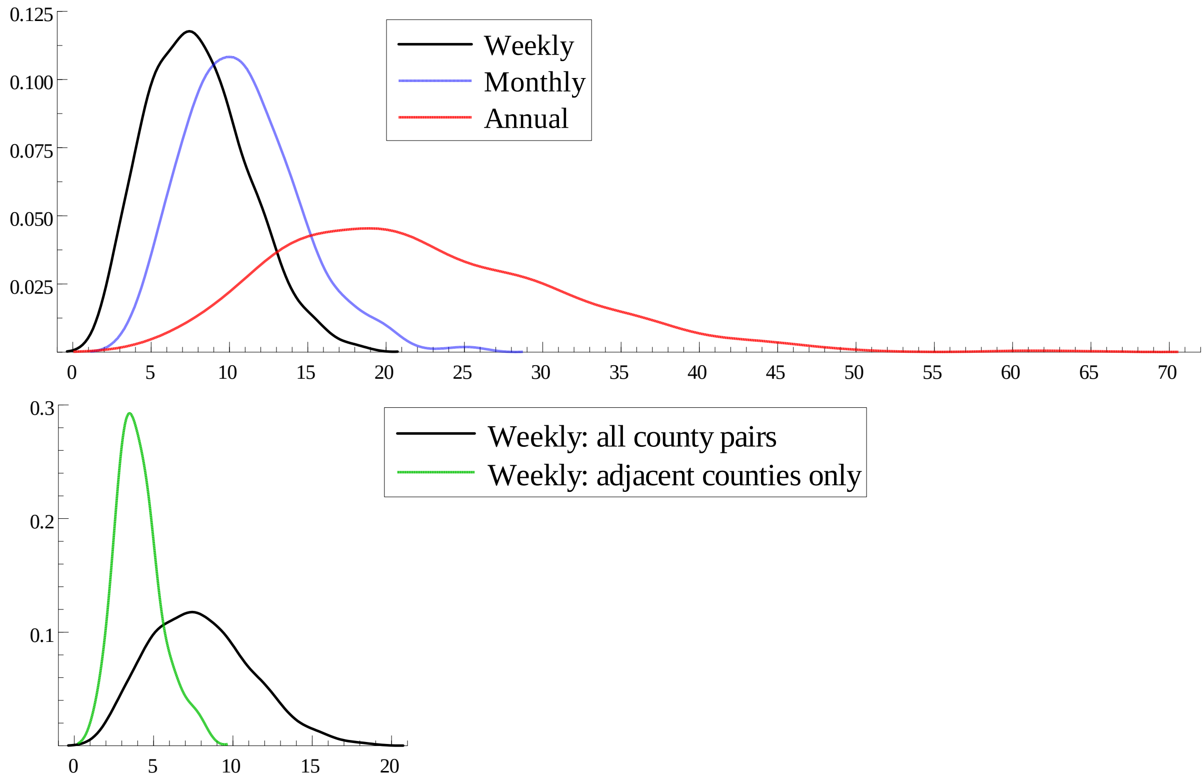
The graph plots the correlations of county prices in each year with prices in the following year (equation 2).

Figure 4: Moran's I Statistics



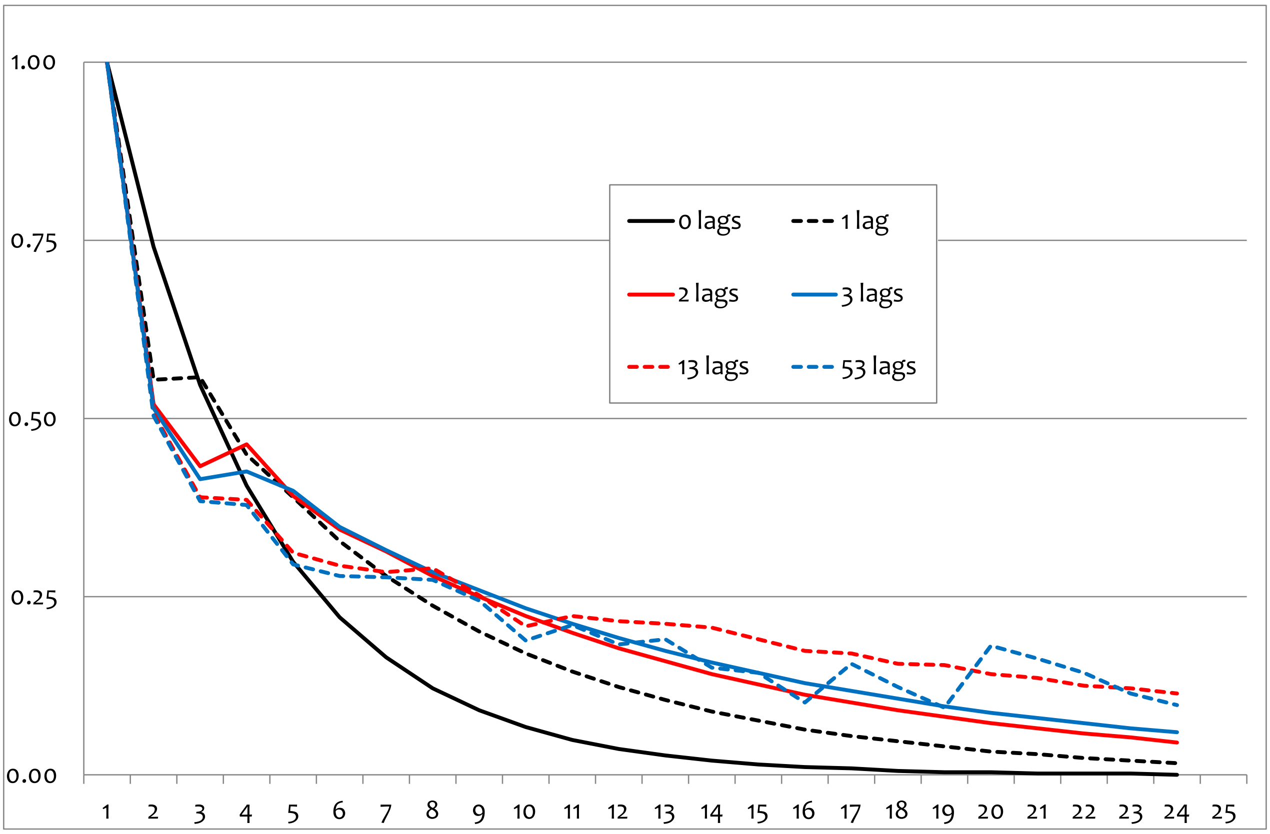
Each point plotted in the figure is a Moran I statistic calculated from a separate cross-section of weekly wheat prices using the formula in equation (3).

Figure 5: Distribution of half lives from models estimated on 1770-1820 data



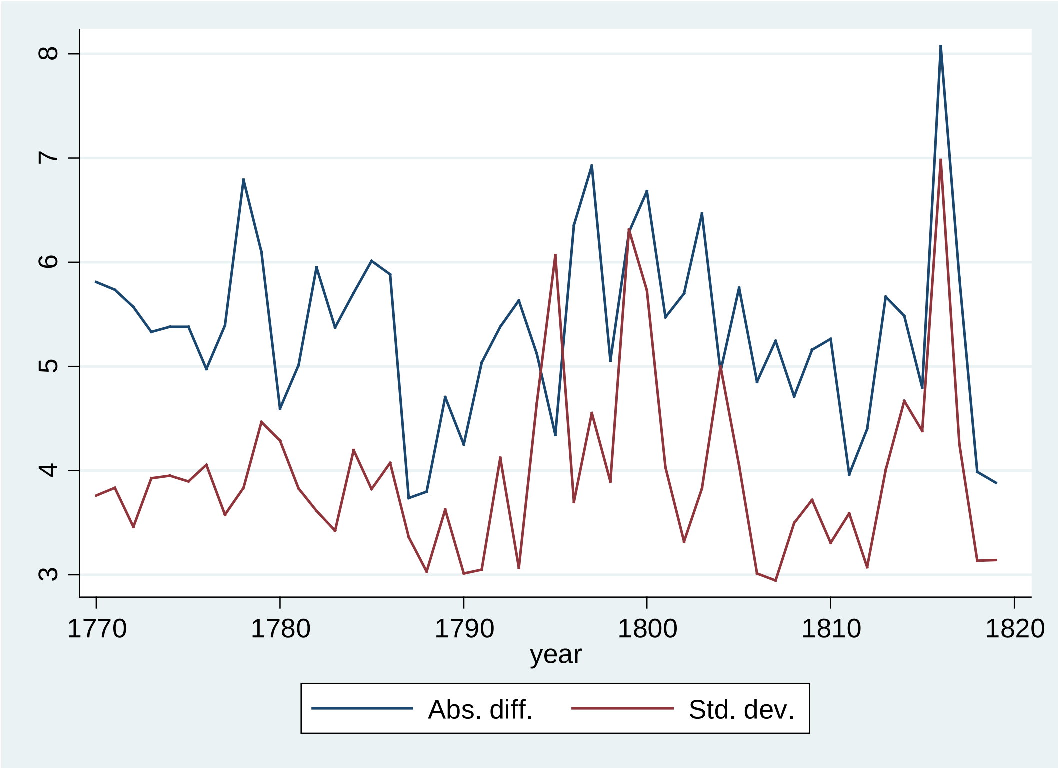
Each distribution in the top panel is based on 780 half lives (slightly fewer for annual data, where some half lives could not be calculated). Each half life is estimated using a model of the form reported in equation (9) using data from the entire period 1770-1820, except where one of the prices is from London, when it is 1770-1793. The bottom panel reproduces the distribution of all 780 half lives from the top panel and compares it to the distribution of the 103 half lives where the counties are adjacent.

Figure 6: Impulse response functions for Bedfordshire-Buckinghamshire



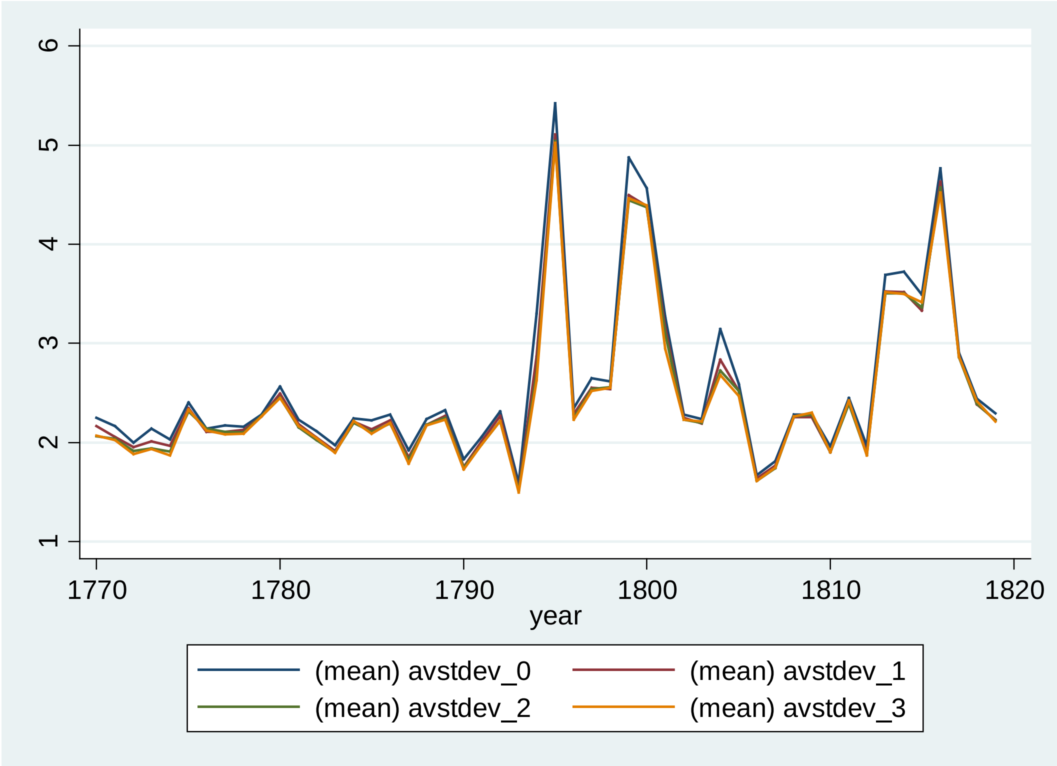
The graph illustrates the speed with which a log-price difference dies away over time (the horizontal axis is measured in weeks). Each impulse response function is estimated using equations (10) to (12). The underlying models are estimated on the full sample of weekly data from 1770-1820 and differ only in the number of lagged dependent variables (the parameter K*).*

Figure 7: Dispersion of prices



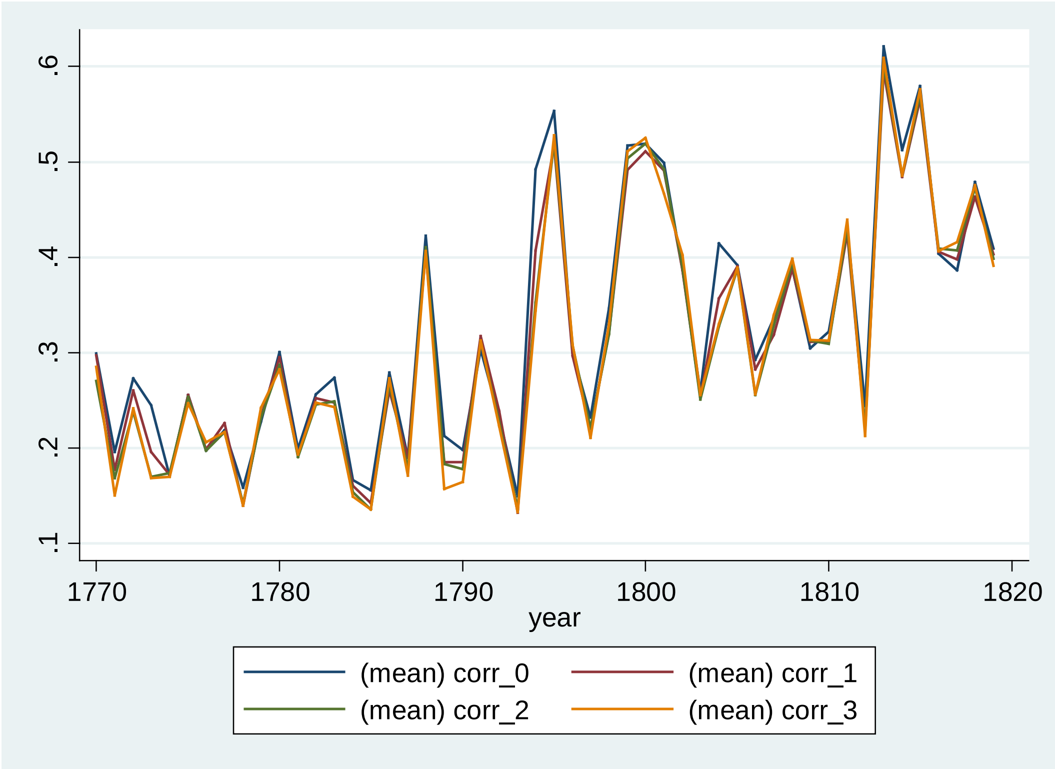
In each year the average for the 93 adjacent-county pairs is plotted of two variables: (i) Abs. diff. is the absolute value of the difference between the mean prices; (ii) Std. dev. is the standard deviation of the price gap. These variables are defined formally in equation (16).

Figure 8: Average magnitude of shocks for each year (per cent)



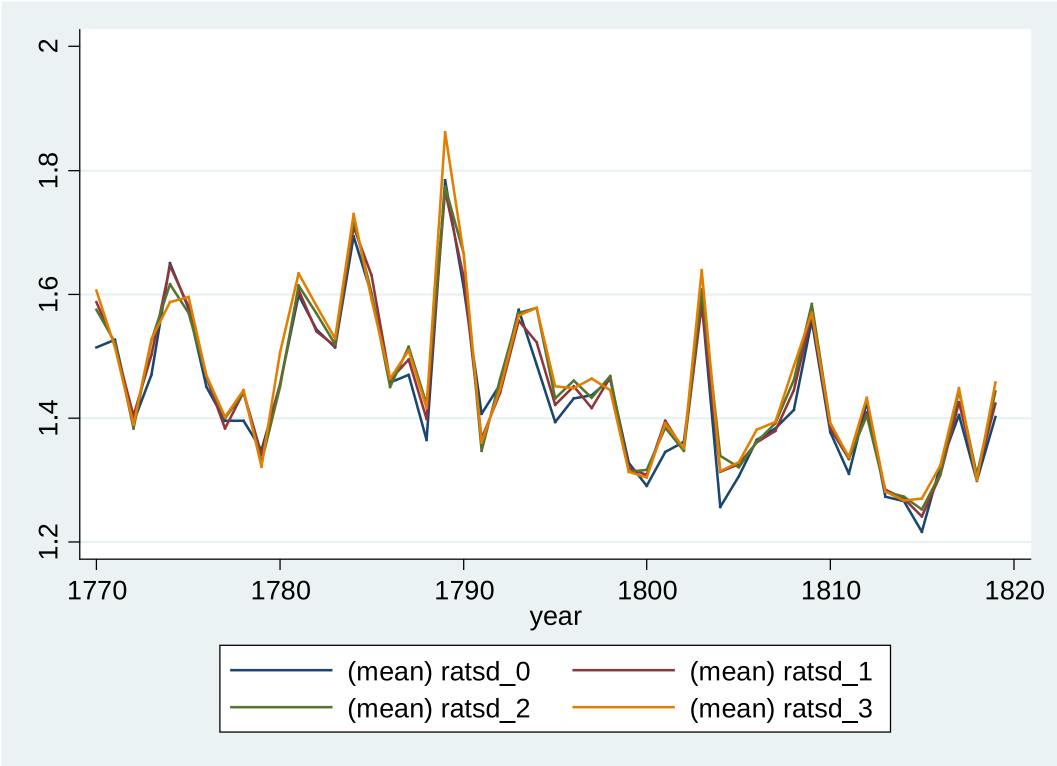
For each year this shows the standard deviation of the shocks (averaged across all 93 adjacent-county pairs) as defined in equation (14) from a model estimated of the form in equation (6). The four series show the estimates from models with 0, 1, 2 or 3 lags in the VECM.

Figure 9: Average correlation of shocks for each year



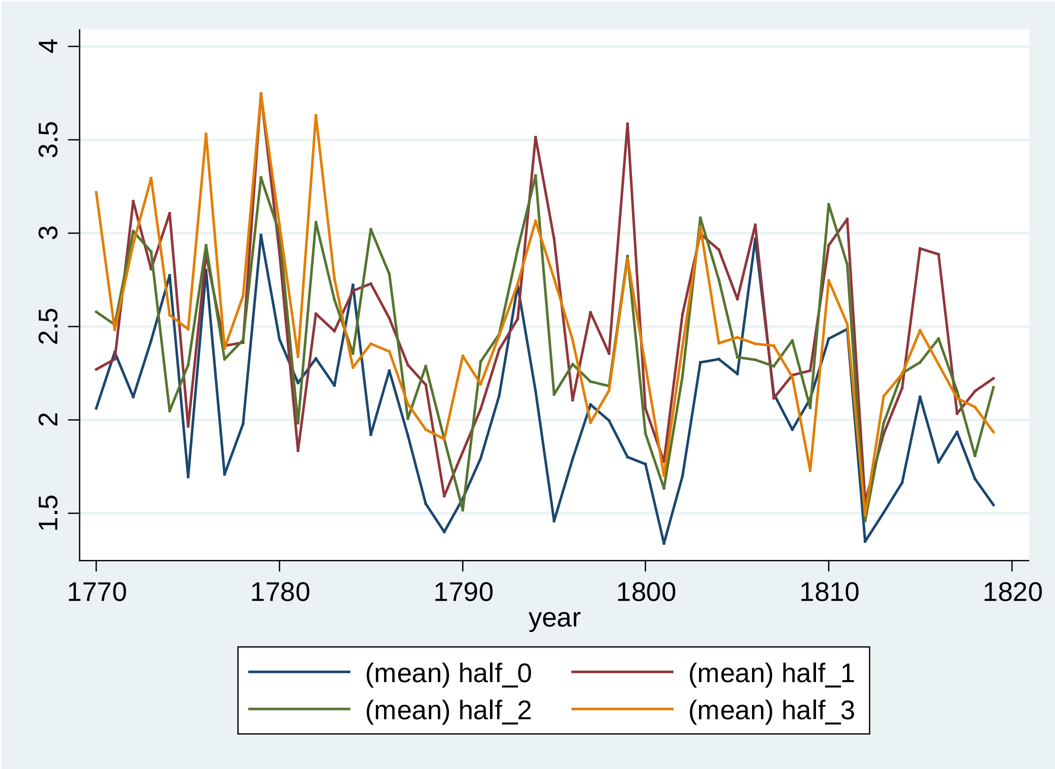
For each year this shows the correlation of the shocks (averaged across all 93 adjacent-county pairs) as defined in equation (15) from a model estimated of the form in equation (6). The four series show the estimates from models with 0, 1, 2 or 3 lags in the VECM.

Figure 10: Average relative size of shocks (ratio)



For each year this shows the ratio of the larger to the smaller standard deviations of the shocks (averaged across all 93 adjacent-county pairs) as defined in equation (16) from a model estimated of the form in equation (6). The four series show the estimates from models with 0, 1, 2 or 3 lags in the VECM.

Figure 11: Average half lives for each year (weeks)



For each year this shows the half life of the response function to a disequilibrium between two pricess (averaged across all 93 adjacent-county pairs). The four series show the estimates from models with 0, 1, 2 or 3 lags in the VECM. When there are no lagged price changes in the half life is the measure defined in equation (8). When there are lags in the first-stage model the half life is calculated using the method described in Appendix 5.2.

Table 1: Effect of including lagged price changes on half-life estimates

Number of lags	0	1	2	3	13	53
Bedfordshire and Buckinghamshire						
Half life based on loadings	2.30	3.18	3.86	4.07	4.66	6.73
Half life from impulse response function	2.33	2.53	1.23	1.14	1.06	1.04
Average for all 780 county pairs						
Half life from impulse response function	7.72	8.50	8.36	8.16	6.66	6.44
Average for the 93 adjacent-county pairs						
Half life from impulse response function	4.09	4.61	4.28	3.71	2.95	2.80

Results are based on regressions for pairs of prices estimated for the whole period 1770-1820 (ignoring issues of parameter stability). The first two rows use weekly data for Bedfordshire and Buckinghamshire as an illustrative example and lags 0, 1 and 2 correspond to equations (10.c), (14) and (15). The half life in row 1 is based upon equation (7) using just the estimates of the parameter α; the half life in row 2 is based upon Figure 6 and linear interpolation is used to see where the impulse response function crosses the line $h = \frac{1}{2}$ (described in Appendix 5.2). Half lives in rows 3 and 4 are calculated analogously to those in row 2. The distribution of the half lives in row 3 is illustrated in Appendix 8.

Table 2: Regression analysis

Dependent variable:	Permanent price gap (LOOP)	Standard deviation of price gap	Components of price dispersion (from VECM with price changes)		
			Average St. Dev.	Ratio St. Dev.	Correlation of disturbances
Roads	0.043 (0.624)	-0.027 (1.310)	-0.030 (2.517)	-0.015 (2.526)	0.002 (0.640)
Canals	0.326 (0.634)	-0.236 (1.967)	-0.171 (2.440)	0.011 (0.305)	0.031 (1.987)
Newspapers	-0.743 (0.421)	-0.735 (1.708)	-0.073 (0.264)	-0.705 (3.874)	0.202 (3.044)
R-squared	0.057	0.356	0.722	0.090	0.340

Results are for six different regressions, each for a panel for fifty harvest years (1770-71 to 1819-20) and 93 adjacent-county pairs: some observations are omitted due to insufficient observations in the first stage so there are only 4,642 observations in total. Each regression has a different explained variable, which has been estimated in the first-stage procedure explained in section 4.2. Other than the Canals, Roads and Newspapers explanatory variables, all regressions include adjacent-county-pair fixed effects and year fixed effects. T-statistics in parentheses are robust to heteroskedasticity and within-group correlation. The R-squared refers to within-group variation.

Table 3: Regressions using different measures of market integration

Lags in 1st-stage VAR:	0	1	2	3
Dependent variable: Half life				
Roads	0.067 (2.316)	0.063 (1.673)	0.039 (1.044)	0.064 (1.508)
Canals	0.307 (2.131)	0.199 (0.852)	0.397 (1.771)	0.281 (1.181)
Newspapers	-0.296 (0.332)	1.132 (1.084)	0.928 (1.051)	-0.243 (0.195)
N × T	4564	4384	4330	4308
R-squared	0.051	0.032	0.030	0.032

Results are for four separate regressions, each one on a panel of annual data for 1771/2 to 1819/20 for each adjacent-county pair. The explained variable is the half life for which each observation is estimated from a separate first-stage VAR on weekly data using the model in equation (6). When there are zero lags in the first-stage VAR the half life is calculated using equation (8); when there are more lags the half life is calculated using the procedure described in Appendix 5.2. T-statistics in parentheses are robust to heteroskedasticity and within-group correlation. The R-squared refers to within-group variation.

Appendix 1: List of Counties

Bedfordshire	Lincolnshire
Berkshire	Middlesex
Buckinghamshire	Monmouthshire
Cambridgeshire	Norfolk
Cheshire	Northamptonshire
Cornwall	Northumberland
Cumberland	Nottinghamshire
Derbyshire	Oxfordshire
Devon	Rutlandshire
Dorset	Shropshire
Durham	Somerset
Essex	Staffordshire
Gloucestershire	Suffolk
Hampshire	Surrey
Herefordshire	Sussex
Hertfordshire	Warwickshire
Huntingdonshire	Westmorland
Kent	Wiltshire
Lancashire	Worcestershire
Leicestershire	Yorkshire

Appendix 2: Correlations of prices over time

In the main text, we asked how the pattern of prices changed from year to year and this was illustrated in Figure 3. This showed that cross-sectional variation in prices in one year was similar to the cross-sectional variation of prices in the following year. This raises the obvious question of how cross-sectional variation of prices changed over longer periods.

As in Figure 3, we start by calculating the within-harvest-year average price for each county

(A2.1)
$$\tilde{p}_{i.y} \equiv \sum_{w \in y} p_{i.w} \Big/ 52$$
$$\text{e.g. } \tilde{p}_{\text{Bedfords.1781-2}} \equiv \frac{p_{\text{Bedfords.1781,week 45}} + p_{\text{Bedfords.1781,week 46}} + \cdots + p_{\text{Bedfords.1782,week 44}}}{52}$$

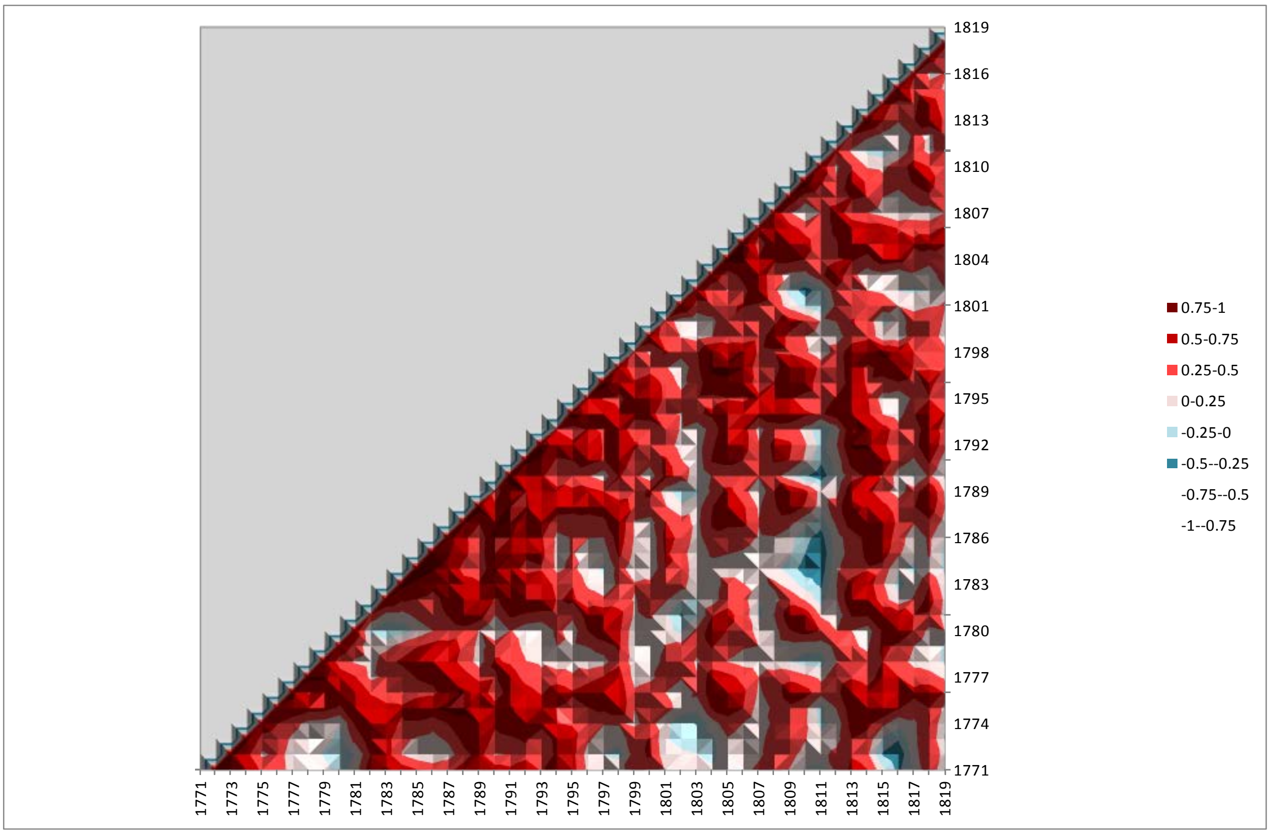
This means that for each harvest year from 1771/2 to 1819/20 we have 43 county prices. For any two harvest years, y and $y+x$, we then can calculate the correlation between the (average) prices for different counties using the conventional correlation coefficient

(A2.2)
$$\text{corr}\left(\tilde{p}_{i.y}, \tilde{p}_{i.y+x}\right) = \frac{\sum_{i=1}^{i=40} \left(\tilde{p}_{i.y} - \bar{\tilde{p}}_{y}\right)\left(\tilde{p}_{i.y+x} - \bar{\tilde{p}}_{y+x}\right)}{\sqrt{\sum_{i=1}^{i=40} \left(\tilde{p}_{i.y} - \bar{\tilde{p}}_{y}\right)^2 \sum_{i=1}^{i=43} \left(\tilde{p}_{i.y+x} - \bar{\tilde{p}}_{y+x}\right)^2}}$$

We illustrate the resulting 946 correlations in the implicitly three-dimensional diagram in Figure A3: the horizontal axis shows the year $y+x$ and the vertical axis year y. The correlation is shown by the colour of the diagram.

For example, if we look at the point corresponding to 1809 on the horizontal axis and 1789 on the vertical axis we see that the area is shaded dark red, so the correlation between county prices in 1789 and 1809 was between 0.50 and 0.75 (in fact it was 0.723). This means that the pattern of prices between different counties in 1789-90 (before the French revolution had really started) was very similar to the pattern of prices in 1809-10 (when France had just defeated Austria for the fifth time and Britain had just embarked on the Peninsular War). In fact, most of the diagram is red or brown, showing that the pattern of prices remained remarkably constant for most of the period.

Figure A2.1: Correlations of cross-sectional price series for all year-pairs



The graph plots the correlation of county prices in each year with all other years.

Appendix 3: Notation

In this section we carefully define our matrix notation. Recall that our most general model in equation (5) is:

$$
\text{(A3.1)}\quad
\begin{aligned}
&\begin{bmatrix}\Delta p_t^i \\ \Delta p_t^j\end{bmatrix} = \begin{bmatrix}\alpha_i \\ \alpha_j\end{bmatrix}\left(p_{t-1}^i - p_{t-1}^j\right) + \begin{bmatrix}\mu_i \\ \mu_j\end{bmatrix} + \begin{bmatrix}\pi_{i.i}^{(1)} & \pi_{i.j}^{(1)} \\ \pi_{j.i}^{(1)} & \pi_{j.j}^{(1)}\end{bmatrix}\begin{bmatrix}\Delta p_{t-1}^i \\ \Delta p_{t-1}^j\end{bmatrix} + \cdots + \begin{bmatrix}\pi_{i.i}^{(K)} & \pi_{i.j}^{(K)} \\ \pi_{j.i}^{(K)} & \pi_{j.j}^{(K)}\end{bmatrix}\begin{bmatrix}\Delta p_{t-K}^i \\ \Delta p_{t-K}^j\end{bmatrix} + \begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix}; \\
&\begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix} \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\omega_{i.i} & \omega_{i.j} \\ \omega_{i.j} & \omega_{j.j}\end{bmatrix}\right); \qquad \alpha_i \le 0; \quad \alpha_j \ge 0; \quad \alpha_j - \alpha_i < 1
\end{aligned}
$$

where $\Delta p_t^i \equiv p_t^i - p_{t-1}^i$ which we refer to as the price change and we re-write this in vector notation as

$$
\text{(A3.2)}\quad \Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\gamma}\mathbf{p}_{t-1} + \boldsymbol{\mu} + \sum_{k=1}^{K} \boldsymbol{\pi}^{(k)} \Delta \mathbf{p}_{t-k} + \boldsymbol{\varepsilon}_t; \quad \boldsymbol{\varepsilon}_t \sim N\left(\mathbf{0}, \boldsymbol{\Omega}\right); \quad \boldsymbol{\gamma} \equiv \begin{bmatrix}1 & -1\end{bmatrix}.
$$

We use the common (although not universal) practice of denoting vectors and matrices with bold type and scalars in light type. The vectors and matrices are defined formally as

$$
\text{(A3.3)}\quad
\begin{aligned}
&\mathbf{p}_t \equiv \begin{bmatrix}p_t^i \\ p_t^j\end{bmatrix}; \quad \Delta \mathbf{p}_t \equiv \begin{bmatrix}\Delta p_{t-s}^i \\ \Delta p_{t-s}^j\end{bmatrix} = \begin{bmatrix}p_t^i - p_{t-1}^i \\ p_t^j - p_{t-1}^j\end{bmatrix}; \quad \boldsymbol{\varepsilon}_t \equiv \begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix} \\
&\boldsymbol{\alpha} \equiv \begin{bmatrix}\alpha_i \\ \alpha_j\end{bmatrix}; \quad \boldsymbol{\gamma} \equiv \begin{bmatrix}1 & -1\end{bmatrix}; \quad \boldsymbol{\mu} \equiv \begin{bmatrix}\mu_i \\ \mu_j\end{bmatrix}; \quad \boldsymbol{\pi}^{(k)} \equiv \begin{bmatrix}\pi_{i.i}^{(k)} & \pi_{i.j}^{(k)} \\ \pi_{j.i}^{(k)} & \pi_{j.j}^{(k)}\end{bmatrix} \\
&\boldsymbol{\Omega} \equiv \begin{bmatrix}\omega_{i.i} & \omega_{i.j} \\ \omega_{i.j} & \omega_{j.j}\end{bmatrix}; \quad \mathbf{I} \equiv \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}; \quad \mathbf{0} \equiv \begin{bmatrix}0 \\ 0\end{bmatrix}
\end{aligned}
$$

There are several points to note about $\boldsymbol{\alpha\gamma}$. First,

$$
\text{(A3.4)}\quad \boldsymbol{\gamma}\mathbf{p}_t \equiv \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}p_t^i \\ p_t^j\end{bmatrix} = p_t^i - p_t^j
$$

which is just the price gap. Second, the matrix $\boldsymbol{\alpha\gamma}$ has imposed three restrictions, two substantive and one an "identifying restriction". To see the first substantive restriction, notice that we could have written

(A3.5) $$\Delta \mathbf{p}_t = \mathbf{\Psi p}_{t-1} + \cdots; \qquad \mathbf{\Psi} \equiv \begin{bmatrix} \psi_{i.i} & \psi_{i.j} \\ \psi_{i.j} & \psi_{j.j} \end{bmatrix}$$

with no restriction on the four parameters. Even with a completely unrestricted version of $\boldsymbol{\gamma}$, which we denote $\boldsymbol{\gamma}^*$, by writing $\mathbf{\Psi} = \boldsymbol{\alpha\gamma}^*$ we have imposed the restriction that $\psi_{i.i}\psi_{j.j} = \psi_{i.j}\psi_{i.j}$: to see this note that

(A3.6) $$\boldsymbol{\alpha\gamma}^* = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \begin{bmatrix} \gamma_i & \gamma_j \end{bmatrix} = \begin{bmatrix} \alpha_i\gamma_i & \alpha_i\gamma_j \\ \alpha_j\gamma_i & \alpha_j\gamma_j \end{bmatrix}.$$

Formally, the restriction consists in restricting the rank of the matrix $\mathbf{\Psi}$ to equal one (instead of two). When the individual price series follow unit root processes this restriction corresponds to saying that there is an equilibrium relationship and this could be tested using Johansen's maximum likelihood procedure. (If there is no unit root, then imposing the restriction still makes economic sense and can be tested using conventional t and F tests).

Despite this restriction, the form of the matrix in (A3.6) is underidentified because we could replace $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ with $2\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}/2$ without making any difference to the product $\boldsymbol{\alpha\gamma}$. This means we need a normalising restriction: although any of the parameters could be normalised, it is convenient here to restrict the model to

(A3.7) $$\boldsymbol{\alpha\gamma}^* = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \begin{bmatrix} 1 & \gamma_j \end{bmatrix}$$

The final restriction that we use in this paper is to restrict $\gamma_j = -1$. Again this can be tested within the Johansen maximum likelihood procedure. In our paper we impose this restriction: where we test the restriction it is virtually never rejected.

Appendix 4: The Constant and Seasonals

In the error-correction model the constant can be either restricted or unrestricted. In our specification we follow papers such as Marks (2010) and place no restriction on the constant terms in the vector $\boldsymbol{\mu}$, but in some published articles the constant appears to be constrained to lie in the cointegrating space (for example, in equation 5A.4 in Ejrnæs' appendix in Persson, 1999, p.157 of Ejrnæs, Persson and Rich, 2008); in others, it appears to be omitted altogether (such as in Buyst, Dercon and Van Campenhout, 2006). In this appendix we clarify what we mean by a restricted constant and discusses the consequences of differing modelling strategies.

The constant plays two rôles in the cointegrating model. For notational simplicity our exposition in this section ignores the lagged-dependent variables and we hence re-write equation (5) as

(A4.1)
$$\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} \mu_i \\ \mu_j \end{bmatrix} + \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix};$$

this model is usually described as having an unrestricted constant. Alternatively it is also possible to restrict the constant to lie in the cointegrating space so that

(A4.2)
$$\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j + \lambda \right) + \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix}.$$

In the restricted version, the equilibrium condition is that $p_{t-1}^i = p_{t-1}^j - \lambda$ (i.e. there is a systematic difference between the price levels). When the market is in equilibrium the expected price change is zero, which means that there is no systematic trend up or down in prices. Notice that this version of the model is the same as the first model with the cross-equation restriction that $\mu_j = \alpha_j \mu_i / \alpha_i$. When the restriction is not imposed, equation (A4.1) can be rewritten as

(A4.3)
$$\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j + \lambda \right) + \begin{bmatrix} \mu_i' \\ \mu_j' \end{bmatrix} + \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix};$$

which emphasises that there can be both a systematic difference between the two prices and a stochastic trend. When the restriction is valid, there is potentially an efficiency gain from imposing the restriction in the estimation; conversely imposing the restriction in the model when it is invalid will bias parameter estimates. It is always possible to test the restriction by using a likelihood ratio test.

In our data, as can be seen from Figure 1, there are often systematic price differences and the overall trend from 1770-1820 is quite small. This has the following consequences, where we summarise analysis that is not reported here or in the main paper.

If we omit the constant altogether and estimate a model for the whole period 1770-1820 then the result is that our other parameter estimates are highly biased, since we are imposing an invalid equilibrium condition that $p_{t-1}^{i} = p_{t-1}^{j}$. Typically the estimated half life is biased up by a factor of as much as two.

If we restrict the constant to lie in the cointegrating space and estimate the model for the whole period 1770-1820, then there is a negligible effect on the estimated half life. The reason for this is that prices at the end of the period are not much higher than at the beginning of the period and so the unrestricted estimate of the drift term is close to zero anyway: the restriction makes little difference

If we restrict the constant to lie in the cointegrating space and estimate the model for a sub-sample of the data, however, then the restriction can have a big impact on the estimated half life. The reason for this is that, over various sub-samples, prices do go up or down by substantial amounts (as can be seen in Figure 1, for example 1803-1812) and therefore it is important to include a stochastic drift term in the model.

In principle we could use a sophisticated process by which the constant was sometimes restricted and sometimes unrestricted, using an appropriate test as the criterion for model selection. However, since we would invariably make some Type I errors, this would involve some invalid restrictions: on the other hand the gain in efficiency from imposing the restriction would be reduced whenever we made a Type II error. This might involve making inappropriate choices (as some tests would with the criteria). For this reason we choose never to restrict the constant.

Throughout the main text of the paper we omit seasonal dummies from our formulae for notational compactness (and when we analyse data at an annual frequency the issue of seasonals does not arise). When we include the seasonals for models estimated on weekly or monthly data the model becomes:

(A4.4)
$$\begin{bmatrix}\Delta p_t^i \\ \Delta p_t^j\end{bmatrix} = \begin{bmatrix}\alpha_i \\ \alpha_j\end{bmatrix}\left(p_{t-1}^i - p_{t-1}^j\right) + \begin{bmatrix}\mu_i \\ \mu_j\end{bmatrix} + \sum_{w=1}^{51}\begin{bmatrix}\delta_w^i \\ \delta_w^j\end{bmatrix} + \begin{bmatrix}\varepsilon_t^i \\ \varepsilon_t^j\end{bmatrix};$$

(If the data were monthly then there would be eleven, rather than 51, seasonals). Restricting the seasonals to the cointegrating space would imply that there were no seasonal effects on expected price changes, but that the equilibrium relationship between the two prices changed over the year, which is a slightly strange assumption and not borne out by the facts (as prices show a seasonal pattern). For this reason we do not restrict the seasonals, either.

Appendix 5: The Half Life

A5.1: The half life when there are no lagged price changes

In this section we discuss several technical issues with the half life. When the VECM model has no lagged price changes, so that it can be written as

(A5.1) $$\Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\gamma}\mathbf{p}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t,$$

then the half life is a sufficient statistic to describe adjustment back to equilibrium from a position of disequilibrium. This is because the decay in the price gap is geometric. As noted in the main text of the paper, the formula for the half life in this instance is

(A5.2) $$HL \equiv \frac{\ln(0.5)}{\ln(1+\alpha_i-\alpha_j)} = \frac{\ln(0.5)}{\ln(1+\boldsymbol{\gamma}\boldsymbol{\alpha})}$$

If we had a sufficiently large sample then we could rely upon a consistency result that

(A5.3) $$\text{plim}\left[\frac{\ln(0.5)}{\ln(1+\boldsymbol{\gamma}\hat{\boldsymbol{\alpha}})}\right] = \frac{\ln(0.5)}{\ln(1+\boldsymbol{\gamma}\mathsf{E}[\hat{\boldsymbol{\alpha}}])} = \frac{\ln(0.5)}{\ln(1+\boldsymbol{\gamma}\boldsymbol{\alpha})} = HL.$$

But some of our results are based on relatively small sub-samples of the data, so we wish to know the properties of

(A5.4) $$\frac{\ln(0.5)}{\ln(1+\boldsymbol{\gamma}\mathsf{E}[\hat{\boldsymbol{\alpha}}])} = \frac{\ln(0.5)}{\ln(1+\hat{\alpha}_i-\hat{\alpha}_j)}$$

in small samples. Most authors simply substitute $\hat{\alpha}_i - \hat{\alpha}_j$ into this formula to estimate the half life. Although the half life is an increasing function of $\alpha_i - \alpha_j$, when this quantity is less than about -0.57 the function is concave; thereafter it is convex. This suggests that the expected value of the half life will not be the same as the half life evaluated at the expected value of the parameters. However, in nearly all cases the standard error of $\hat{\alpha}_i - \hat{\alpha}_j$ is sufficiently small that it makes no difference: the reason for this is that $\hat{\alpha}_i$ and $\hat{\alpha}_j$ are highly negatively correlated and the variance of $\alpha_i - \alpha_j$ is correspondingly quite low.

As a further check, we tried a Monte Carlo procedure to see if the non-linearity made any difference. To do this we assumed that the disturbances had a Normal distribution (which is only approximately correct), so that

(A5.5) $$\widehat{\alpha_i - \alpha_j} \sim \mathsf{N}\left(\alpha_i - \alpha_j, \mathsf{var}\left(\widehat{\alpha_i - \alpha_j}\right)\right).$$

Note that

(A5.6) $$\mathsf{var}\left(\widehat{\alpha_i - \alpha_j}\right) = \mathsf{var}\left(\hat{\alpha}_i\right) + \mathsf{var}\left(\hat{\alpha}_j\right) - 2\,\mathsf{cov}\left(\hat{\alpha}_i, \hat{\alpha}_j\right).$$

Using this as a basis, we simulated 10,000 values $\alpha_i - \alpha_j$ from a Normal distribution $N\left(\widehat{\alpha_i - \alpha_j}, \mathrm{var}\left(\widehat{\alpha_i - \alpha_j}\right)\right)$ and calculated the corresponding 10,000 half lives (in a very small number of cases the draw of $\alpha_i - \alpha_j$ was negative and these were discarded). We then averaged the 10,000 replications and compared the mean to the conventionally calculated half life. We found that in nearly all cases the standard error of $\widehat{\alpha_i - \alpha_j}$ was sufficiently small that it made no real difference which method we used.

A5.2: The half life when there are lagged price changes

In the general case there are lagged price changes and the model can be written as

(A5.7) $$\Delta \mathbf{p}_t = \boldsymbol{\alpha\gamma}\mathbf{p}_{t-1} + \boldsymbol{\mu} + \sum_{k=1}^{K} \boldsymbol{\pi}^{(k)} \Delta \mathbf{p}_{t-k} + \boldsymbol{\varepsilon}_t.$$

In this case there is no single measure which summarises the speed of adjustment. To understand the difference between this situation and that in (A5.1), consider the hypothetical possibility that prices in period $t-1$ were $p_{t-1}^i - p_{t-1}^j = 0.1$ so that price *i* were approximately ten per cent higher than price *j*. Conceptually, we can distinguish two simple processes that could have resulted in this price gap: either (i) $p_{t-2}^i - p_{t-2}^j = 0$ and there was a shock to the prices in period $t-1$; or (ii) there was no shock in period $t-1$ but $p_{t-1}^i - p_{t-1}^j > 0.1$ and the price gap existed in $t-1$ because prices had not yet fully adjusted back to equilibrium after a shock in period $t-2$ or earlier. With the model in (A5.1), the price adjustment in period t would be identical: it is as if the process generating prices had “forgotten” how the disequilibrium had arisen. With the more general model, the price behaviour in period t would depend upon whether the price gap had arisen from situation (i) or situation (ii). Since the adjustment in period t depends upon the earlier behaviour of prices there cannot be a single measure of the speed of adjustment.

Despite this we wish to summarise the speed of adjustment, even if our measure be imperfect. The method we choose is to plot an impulse response function like that in Figure 6 and then see where this curve crosses the horizontal line $y = 0.5$.

This raises the further issue of how to plot the impulse response function and this is complicated because the impulse response function to a shock to price *i* (i.e. due to a shock ε_t^i) may differ from a shock to price *j* (i.e. due to a shock ε_t^j). In the three-price context of New York, London and Copenhagen, Ejrnæs, Persson and Rich (2006) illustrate impulse reponse functions to shocks in all three cities to all three price series. In a two-price context, these correspond to the effect on prices of shocks of the form

(A5.8) $$\text{either } \boldsymbol{\varepsilon}_t \equiv \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ or } \boldsymbol{\varepsilon}_t \equiv \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

i.e. a shock to one price with no effect on the other. The problem with this is that it is rare for price shocks to occur in isolation and we know that ε_t^i and ε_t^j are correlated (as we see in Figure 9). Fortunately, although the response of the two individual prices depends upon the nature of the shock, Pesaran and Shin (1996) show that the speed of adjustment towards equilibrium is the same regardless of which random shock causes the initial price difference. They suggest a method for calculating the impulse response function as follows. First, re-write the VECM of equation (A5.7) in the form of a VAR:

(A5.9) $$\begin{aligned} &\mathbf{p}_t = \boldsymbol{\Phi}_1\mathbf{p}_{t-1} + \boldsymbol{\Phi}_2\mathbf{p}_{t-2} + \cdots + \boldsymbol{\Phi}_{K+1}\mathbf{p}_{t-K-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \\ &\boldsymbol{\Phi}_1 \equiv \left(\mathbf{I} + \boldsymbol{\alpha\gamma} + \boldsymbol{\pi}_1\right); \quad \boldsymbol{\Phi}_2 \equiv \left(\boldsymbol{\pi}_2 - \boldsymbol{\pi}_1\right); \ldots \quad \boldsymbol{\Phi}_K \equiv \left(\boldsymbol{\pi}_K - \boldsymbol{\pi}_{K-1}\right); \quad \boldsymbol{\Phi}_{K+1} \equiv -\boldsymbol{\pi}_K \end{aligned}$$

From a hypothetical position of equilibrium $p_{t-2}^i - p_{t-2}^j = 0$, the effect of a shock in period *s* relative to what would have happened if there had been no shock is then calculated iteratively via

(A5.10) $$\begin{aligned} \mathbf{p}_0 &= \boldsymbol{\varepsilon}_0 \\ \mathbf{p}_1 &= \boldsymbol{\Phi}_1\boldsymbol{\varepsilon}_0 \\ \mathbf{p}_2 &= \boldsymbol{\Phi}_1\boldsymbol{\Phi}_1\boldsymbol{\varepsilon}_0 + \boldsymbol{\Phi}_2\boldsymbol{\varepsilon}_0 \\ \mathbf{p}_3 &= \boldsymbol{\Phi}_1\left(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1\boldsymbol{\varepsilon}_0 + \boldsymbol{\Phi}_2\boldsymbol{\varepsilon}_0\right) + \boldsymbol{\Phi}_2\boldsymbol{\Phi}_1\boldsymbol{\varepsilon}_0 + \boldsymbol{\Phi}_3\boldsymbol{\varepsilon}_0 \\ &\vdots \end{aligned}$$

Pesaran and Shin (1996) suggest that a potential and natural shock to use is

(A5.11) $$\sqrt{\boldsymbol{\gamma}\boldsymbol{\Omega}\boldsymbol{\gamma}'} = \sqrt{\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\omega_{i.i} & \omega_{i.j} \\ \omega_{i.j} & \omega_{j.j}\end{bmatrix}\begin{bmatrix}1 \\ -1\end{bmatrix}} = \sqrt{\omega_{i.i} + \omega_{j.j} - 2\omega_{i.j}}$$

and they then calculate the effect of this shock. To implement this, they define

(A5.12) $$\begin{aligned} &\mathbf{B}_s \equiv \mathbf{0} \quad \text{for} \quad s < 0 \\ &\mathbf{B}_0 \equiv \mathbf{I} \\ &\mathbf{B}_s \equiv \sum_{k=1}^{K} \Phi_k \mathbf{B}_{s-k} \quad \text{for} \quad s > 0 \end{aligned}$$

in which case the impulse response function (normalised by adjusting for the variance of the original shock) is

(A5.13) $$i(s) = \sqrt{\frac{\gamma \mathbf{B}_s \Omega \mathbf{B}'_s \gamma'}{\gamma \Omega \gamma'}}$$

which is the formula used to derive the functions in Figure 6.

Appendix 6: Decomposing the RMS price difference when there are lagged dependent variables.

In this section we derive the decomposition of equation (9) more formally. Recall the general VECM from equation (6)

$$\Delta \mathbf{p}_t = \alpha\gamma\mathbf{p}_{t-1} + \mu + \sum_{k=1}^{K} \pi^{(k)} \Delta \mathbf{p}_{t-k} + \varepsilon_t \tag{A6.1}$$

which, by adding $\mathbf{p}_{t-1}$ to both sides, can be re-written as

$$\mathbf{p}_t = \left(\mathbf{I} + \alpha\gamma\right)\mathbf{p}_{t-1} + \mu + \sum_{k=1}^{K} \pi^{(k)} \Delta \mathbf{p}_{t-k} + \varepsilon_t . \tag{A6.2}$$

Multiplying by gg gives

$$\gamma\mathbf{p}_t = \gamma\left(\mathbf{I} + \alpha\gamma\right)\mathbf{p}_{t-1} + \gamma\mu + \sum_{k=1}^{K} \gamma\pi^{(k)} \Delta \mathbf{p}_{t-k} + \gamma\varepsilon_t . \tag{A6.3}$$

Consider first the simplest case where there are no lagged differences so that $\pi^{(k)} = \mathbf{0}$. Note also that $\gamma\left(\mathbf{I} + \alpha\gamma\right) = \gamma + \gamma\alpha\gamma = \left(1 + \gamma\alpha\right)\gamma$, so that the simple case becomes

$$\gamma\mathbf{p}_t = \left(1 + \alpha\gamma\right)\gamma\mathbf{p}_{t-1} + \gamma\mu + \gamma\varepsilon_t \tag{A6.4}$$

or, since the formula consists of scalars,

$$\left(p_t^i - p_t^j\right) = \left(1 + \alpha_i - \alpha_j\right)\left(p_{t-1}^i - p_{t-1}^j\right) + \left(\mu_i + \mu_j\right) + \left(\varepsilon_t^i + \varepsilon_t^j\right). \tag{A6.5}$$

Squaring this formula and taking expectations yields (in matrix and scalar notation respectively)

$$\begin{aligned} \mathsf{E}\left[\gamma\mathbf{p}_t\mathbf{p}_t'\gamma'\right] &= \left(1 + \gamma\alpha\right)^2 \mathsf{E}\left[\gamma\mathbf{p}_{t-1}\mathbf{p}_{t-1}'\gamma'\right] + \gamma\mu\mu'\gamma' + \gamma\Omega\gamma' \\ \mathsf{E}\left[\left(p_t^i - p_t^j\right)^2\right] &= \left(1 + \alpha_i - \alpha_j\right)^2 \mathsf{E}\left[\left(p_{t-1}^i - p_{t-1}^j\right)^2\right] + \left(\mu_i - \mu_j\right)^2 + \omega_{i.i} + \omega_{j.j} + 2\omega_{i.j} \end{aligned}, \tag{A6.6}$$

which is the same as equation (10) as discussed in the text: the squared price difference *this period* depends upon the shocks *this period*; the constants; and the price dispersion *in the previous period* multiplied by a term showing the speed of adjustment.

The more general case is messier but has the same underlying intuition. From equation (A6.3) we obtain

$$\begin{aligned}
\mathrm{E}\left[\gamma \mathbf{p}_t \mathbf{p}_t' \gamma'\right] &= \left(1+\gamma\alpha\right)^2 \mathrm{E}\left[\gamma \mathbf{p}_{t-1} \mathbf{p}_{t-1}' \gamma'\right] + \gamma\mu\mu'\gamma' + \gamma\Omega\gamma' \\
&\quad + \mathrm{E}\left[\sum_{k=1}^{K} \gamma\pi^{(k)} \Delta\mathbf{p}_{t-k} \mathbf{p}_{t-1}' \gamma' + \gamma \mathbf{p}_{t-1} \sum_{l=1}^{K} \Delta\mathbf{p}_{t-l}' \pi^{(l)\prime} \gamma'\right]. \\
&\quad + \mathrm{E}\left[\sum_{l=1}^{K} \sum_{k=1}^{K} \gamma\pi^{(k)} \Delta\mathbf{p}_{t-k} \Delta\mathbf{p}_{t-l}' \pi^{(l)\prime} \gamma'\right]
\end{aligned} \tag{A6.7}$$

The first row of the right-hand side of this formula is the same as (A6.6): the difference lies in the complicated set of terms in the second and third rows. What these terms denote are the adjustments to price dispersion in previous periods. Recall that the presence of lagged price changes corresponds to a complicated adjustment process to price dispersion in previous periods. Therefore, to describe perfectly the adjustment back towards equilibrium requires a full knowledge of the behaviour of prices over the previous *k* periods. In our analysis in section 5 we summarise this with a single statistic, namely the half life, calculated as described in the text.

Appendix 7: Within-period price adjustment and the interpretation of the parameter $\omega_{i,j}$

Our data in this paper are weekly data and many of the markets we analysed opened only one or two days of the week. This means that there is no or very little temporal aggregation of the form discussed by Taylor (2001). However, even without temporal aggregation, infrequent sampling affects our interpretation of some of the parameters. In this appendix we consider the effect on the parameters in which we are most interested in this paper. For expositional purposes we consider the simplest version of our model, namely

(A7.1)
$$\begin{bmatrix} \Delta p_t^i \\ \Delta p_t^j \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix} \left(p_{t-1}^i - p_{t-1}^j \right) + \begin{bmatrix} \mu_i \\ \mu_j \end{bmatrix} + \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^j \end{bmatrix};$$

which can more conveniently be written as

(A7.2)
$$\mathbf{p}_t = \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \mathbf{p}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t.$$

To separate the effects of infrequent sampling and time aggregation, let us assume that markets traded twice per week and suppose (counter-factually) that we observed end-of-week prices: this would mean that we would observe prices only for $t = \left\{ 2, 4, 6, \ldots \right\}$. Then the relationship between one end-of-week price and the previous end-of-week price would be

(A7.3)
$$\begin{aligned}
\mathbf{p}_t &= \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \underbrace{\left\{ \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \mathbf{p}_{t-2} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_{t-1} \right\}}_{=\mathbf{p}_{t-1}} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \\
&= \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right)^2 \mathbf{p}_{t-2} + \left(2\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \boldsymbol{\mu} + \left\{ \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t \right\} \\
\mathbf{p}_t - \mathbf{p}_{t-2} &= \left(2\boldsymbol{\alpha} + \boldsymbol{\alpha\gamma\alpha} \right) \boldsymbol{\gamma} \mathbf{p}_{t-2} + \left(2\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \boldsymbol{\mu} + \left\{ \left(\mathbf{I} + \boldsymbol{\alpha\gamma} \right) \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t \right\} \\
\left(2\boldsymbol{\alpha} + \boldsymbol{\alpha\gamma\alpha} \right) &= \begin{bmatrix} \alpha_i \left(2 + \alpha_i - \alpha_j \right) \\ \alpha_j \left(2 + \alpha_i - \alpha_j \right) \end{bmatrix}
\end{aligned}$$

If we change the dating convention this can be re-written as

(A7.4)
$$\Delta \mathbf{p}_w \equiv \mathbf{p}_w - \mathbf{p}_{w-1} = \boldsymbol{\alpha}^* \boldsymbol{\gamma} \mathbf{p}_{w-1} + \boldsymbol{\mu}^* + \boldsymbol{\varepsilon}_w^*$$

where the stars indicates the parameters from the weekly data. We can now ask what parameters we shall estimate. The loadings will be

(A7.5) $$\boldsymbol{\alpha}^* \equiv \left(2\boldsymbol{\alpha} + \boldsymbol{\alpha\gamma\alpha}\right) = \begin{bmatrix} \alpha_i\left(2+\alpha_i-\alpha_j\right) \\ \alpha_j\left(2+\alpha_i-\alpha_j\right) \end{bmatrix}$$

Our estimated speed of adjustment will be based on

(A7.6) $$1+\alpha_i^*-\alpha_j^* = 1+\alpha_i\left(2+\alpha_i-\alpha_j\right)-\alpha_j\left(2+\alpha_i-\alpha_j\right)=\left(1+\alpha_i-\alpha_j\right)^2$$

which confirms that we shall estimate exactly the same speed of adjustment: the difference between $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ is entirely due to the different units of measurement (weekly versus half-weekly respectively).

When we turn to the disturbances, whose covariance matrix can be derived as follows:

(A7.7) $$\begin{aligned} \mathsf{var}\left[\boldsymbol{\varepsilon}_w^*\right] &= \mathsf{E}\left[\left\{\left(\mathbf{I}+\boldsymbol{\alpha\gamma}\right)\boldsymbol{\varepsilon}_{t-1}+\boldsymbol{\varepsilon}_t\right\}\left\{\left(\mathbf{I}+\boldsymbol{\alpha\gamma}\right)\boldsymbol{\varepsilon}_{t-1}+\boldsymbol{\varepsilon}_t\right\}'\right] \\ &= \left(\mathbf{I}+\boldsymbol{\alpha\gamma}\right)\boldsymbol{\Omega}\left(\mathbf{I}+\boldsymbol{\alpha\gamma}\right)'+\boldsymbol{\Omega} \\ &= 2\boldsymbol{\Omega}+\left\{\boldsymbol{\alpha\gamma\Omega}+\boldsymbol{\Omega\gamma'\alpha'}+\boldsymbol{\alpha\gamma\Omega\gamma'\alpha'}\right\} \\ &= 2\begin{bmatrix} \omega_{i.i} & \omega_{i.j} \\ \omega_{i.j} & \omega_{j.j} \end{bmatrix} \\ &\quad + \begin{bmatrix} 2\alpha_i\left(\omega_{i.i}-\omega_{i.j}\right) & \alpha_i\left(\omega_{i.j}-\omega_{j.j}\right)+\alpha_j\left(\omega_{i.i}-\omega_{i.j}\right) \\ \alpha_i\left(\omega_{i.j}-\omega_{j.j}\right)+\alpha_j\left(\omega_{i.i}-\omega_{i.j}\right) & 2\alpha_j\left(\omega_{i.j}-\omega_{j.j}\right) \end{bmatrix} \\ &\quad + \left(\omega_{i.i}-2\omega_{i.j}+\omega_{j.j}\right)\begin{bmatrix} \alpha_i^2 & \alpha_i\alpha_j \\ \alpha_i\alpha_j & \alpha_j^2 \end{bmatrix} \end{aligned}$$

The covariance between the disturbances from weekly data consists of the actual covariance of the underlying (half-weekly) disturbances (the parameter $\omega_{i.j}$) and the adjustment which takes place within the week, which is

(A7.8) $$\alpha_i\left(\omega_{i.j}-\omega_{j.j}\right)+\alpha_j\left(\omega_{i.i}-\omega_{i.j}\right)+\alpha_i\alpha_j\left(\omega_{i.i}-2\omega_{i.j}+\omega_{j.j}\right)$$

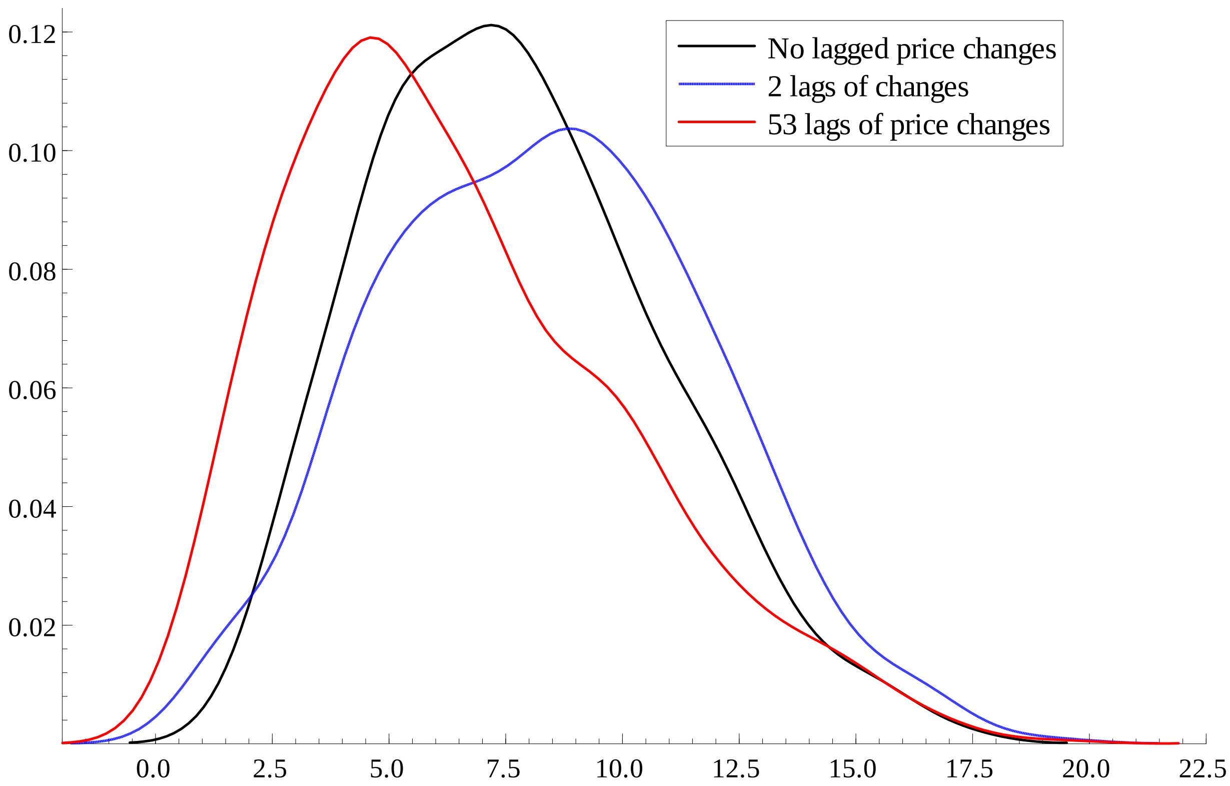
Appendix 8: Additional Tables and Figures

Table A8.1: Summary of half lives estimated for 1770-1820

	All county pairs			Adjacent-county pairs		
frequency of data	weekly	monthly	annual	weekly	monthly	annual
mean	8.0	10.8	22.1	4.1	6.9	22.9
median	7.7	10.4	20.7	3.7	6.4	21.2
st.dev.	3.2	3.7	9.0	1.4	2.2	10.9
minimum	1.5	3.3	5.5	1.5	3.3	5.5
maximum	18.7	25.8	65.0	8.0	14.0	60.9

This table describes the same econometric analysis as that illustrated in Figure 5. The first three columns summarise the distribution of 780 half lives (slightly fewer for annual data, where some half lives could not be calculated). Each half life is estimated from a regression of the form reported in equation (9) using data from the entire period 1770-1820, except where one of the prices is from London, when it is 1770-1793. The final three columns report analogous statistics for the 103 pairs where the counties are adjacent.

Figure A8.2: Distributions of half life estimates



These half lives correspond to the estimation in section 4.1, where a single model is estimated for the whole period (ignoring issues of parameter instability over the period).

Each of these distributions summarises the half lives from 780 regressions, each of which is estimated on weekly data for a county pair for the entire period 1770-1820. Every county pair is estimated, not just adjacent-county pairs.

The only difference between the distributions is the number of lagged price changes used in the regression. These distributions are based on the same information as the third row of Table 1. Note that all the half lives were positive but an artefact of the kernel smoothing method used to estimate the density was that the curves appear to extend to the left of the origin.

Table A8.3: Regressions using different measures of market integration

Lags in 1st-stage VAR:	0	1	2	3
Dependent variable: Average standard deviation of disturbances				
Roads	-0.030 (2.517)	-0.024 (1.941)	-0.025 (1.965)	-0.026 (2.056)
Canals	-0.171 (2.440)	-0.196 (2.641)	-0.205 (2.720)	-0.204 (2.687)
Newspapers	-0.073 (0.264)	-0.034 (-0.117)	-0.124 (0.418)	-0.111 (0.366)
N × T	4642	4642	4642	4642
R-squared	0.722	0.688	0.680	0.674
Dependent variable: Ratio of standard deviations of disturbances				
Roads	-0.015 (2.526)	-0.014 (2.250)	-0.013 (2.079)	-0.013 (2.087)
Canals	0.011 (0.305)	0.015 (0.465)	0.005 (0.142)	0.005 (0.132)
Newspapers	-0.705 (3.874)	-0.709 (3.760)	-0.674 (3.524)	-0.601 (3.216)
N × T	4642	4642	4642	4642
R-squared	0.090	0.085	0.084	0.086
Dependent variable: Correlation of disturbances				
Roads	0.002 (0.640)	0.004 (1.236)	0.003 (1.039)	0.003 (1.122)
Canals	0.032 (1.987)	0.029 (1.717)	0.032 (1.851)	0.036 (2.060)
Newspapers	0.202 (3.044)	0.213 (3.320)	0.188 (2.939)	0.187 (2.953)
N × T	4642	4642	4642	4642
R-squared	0.340	0.316	0.313	0.302
Dependent variable: Half life				
Roads	0.067 (2.316)	0.063 (1.673)	0.039 (1.044)	0.064 (1.508)
Canals	0.307 (2.131)	0.199 (0.852)	0.397 (1.771)	0.281 (1.181)
Newspapers	-0.296 (0.332)	1.132 (1.084)	0.928 (1.051)	-0.243 (0.195)
N × T	4564	4384	4330	4308
R-squared	0.051	0.032	0.030	0.032

Results are for sixteen separate regressions. The explained variables are themselves estimated from regressions on weekly data for each county pair: the column headings refer to the number of lags in the first-stage time-series regressions.