NHH

INSTITUTT FOR SAMFUNNSØKONOMI

DEPARTMENT OF ECONOMICS

SAM 6 2015

ISSN: 0804-6824

March 2015

Discussion paper

Variations in the price and quality of English grain, 1750-1914: quantitative evidence and empirical implications.

BY

Liam Brunt AND **Edmund Cannon**

This series consists of papers with limited circulation, intended to stimulate discussion.

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS

Variations in the price and quality of English grain, 1750-1914: quantitative evidence and empirical implications.

Liam Brunt and Edmund Cannon

Abstract[1]

Interpretation of historic grain price data may be hazardous owing to systematic grain quality variation – both cross sectionally and over varying time horizons (intra-year, inter-year, long run). We use the English wheat market, 1750-1914, as an example to quantify this issue. First, we show that bushel weight approximates grain quality. Then we show that cross sectional and intra-year variation are substantial and problematic, generating erroneous inference regarding market integration. Long run variation is significant, due to sharply declining international quality differentials, and this impacts estimated cost of living changes. By contrast, inter-year variation is smaller and controlled for more easily.

JEL Codes: N01, N50, Q13.

Keywords: Grain quality, markets, cost of living.

[1] Brunt: Department of Economics, NHH – Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway; liam.brunt@nhh.no. Cannon: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, England; edmund.cannon@bris.ac.uk. We thank Paul Poulton at the Rothamsted Experimental Station for providing us with the data for the Rothamsted seed trials. We thank conference participants at the Cliometrics Society Conference at Clemson University, and the Economic History Society Conference at Warwick University – especially Paul Sharp – for helpful comments; we also thank Paul Rhode, two anonymous referees and the editor. We take responsibility for any remaining errors.

1. Introduction

We describe and quantify grain quality variation and discuss the consequences for interpreting price and quantity data. Grain prices constitute a widely available historical source, used to discuss such varied questions as cognitive ability (Baten, Crayen and Voth, 2014) and interest rates (McCloskey and Nash, 1984), as well as more obvious questions such as standard of living (Phelps Brown and Hopkins, 1956), market integration (Shiue and Keller, 2007) and elasticity of demand (Fogel, 1992). Grain remains the most important product in the world economy – without it, we would be unable to meet global population nutrition requirements. This was more obviously true in earlier generations, when grain occupied a large share of household expenditure and calorific intake in even rich economies (Feinstein, 1998, 365), and when the possibility of famine still haunted some European economies (the widespread "year without a summer" of 1816, Irish and Scottish famines in 1847, Finland in 1866-8 and Russia in 1891-2). We consider quality variation of several grains but concentrate on wheat, the primary Western European food grain for the last two hundred years (Collins, 1993).

Quality variation matters for both market participants and economic historians interpreting data. We focus mainly on interpretation, to gauge the sensitivity of modern economic studies to quality variation. Nearly all pre-1914 price and quantity data are characterized by unobserved quality variation, effectively a form of measurement error. The relevance of this to statistical analysis depends partly upon measurement error magnitude, which we quantify, and partly upon data usage.

English grain merchants could determine quality by inspection, both at time of purchase and at delivery. Velkar (2012) documents how increasing mid-nineteenth century international trade resulted in creation of organizations such as the London Corn Trade Association and the Chicago Board of Trade to enable international grain traders to establish quality remotely. Reliable long distance transmission of quality information was crucial, given increasing trade distances and increasing importance for modern milling techniques of knowing wheat quality. However, wheat quality information remained problematic for some imports even in the late nineteenth century – notably Indian, as evidenced by the Secretary of State of India's enquiry of 1885-90 (BPP, 1894, c.7440 and c.7441).

Quality variation is frequently acknowledged in economic history but rarely addressed. Olmstead and Rhode (2003) offer a study of post-1920 US cotton quality improvements. Cotton has an obvious quality metric – staple length (although this is not the only aspect of cotton quality, as they discuss). But they really focus on governmental and economic institutions facilitating rising quality, very different to our focus on actual quality measurement and implications of quality variation for price variation. Olmstead and Rhode (2002) offers an exhaustive study of US wheat seed variety ("cultivar") changes, 1840-1940. The thrust is that changing cultivars both enabled wheat production to spread to harsher US climates and maintained yields in the face of crop pests. Again, this is rather different to our focus on how cultivar affected yield quality, rather than quantity.

Brunt and Cannon (2013) argue that substantial and systematic wheat quality variation makes price movement interpretation difficult – especially pre-1914, when grain quality was not closely regulated. This is important, given the many historical wheat price series available and the many studies basing historical inference on them (examples noted above). We should quantify the four quality variation features that Brunt and Cannon highlight: variation across localities; variation through the year; variation from one year to the next; and trends in quality. The following sections address all these issues. We begin in the next section by defining more carefully what we mean by grain quality and describing how it might vary.

Overall, we find substantial spatial and long run quality variation. There was little *systematic* intra-year quality variation for wheat (rather more for barley and oats); but *random* variation is highly problematic. There was measurable, but modest, inter-year variation in quality.

2. Grain quality

We focus on England, but begin with a quality benchmark based on historical data where quality is best measured – late nineteenth century USA. The Chicago Board of Trade developed a wheat grading system to facilitate exports, culminating in the 1916 US Grain Standards Act (Hill, 1990). Long distance trade (Chicago to New York, thence Europe) necessitated explicit grain quality measures because homogeneity and transparency facilitate trade between distant markets (Henry and Kettlewell, 1996). Each wheat type was subdivided into six grades (Grade 1 at the top, down to Grade 5, to "Sample Grade" at the bottom). Grades were based on bushel weight, moisture content, percentage of damaged kernels, purity, cleanliness and condition (Ball *et al*., 1921).

US price data for this period are for constant quality, and are sometimes available for several grades. Despite some variation, Grade 2 wheat traded at about a 5% discount to Grade 1 wheat, and a 12% premium to Grade 3. Information on proportions of each grade shipped from

Chicago to New York show it varied greatly year-on-year: only 1% of wheat shipped in 1879 was Grade 1, compared to 11% in 1878 (Chicago Board of Trade annual reports – details in appendix A1). Using prices and proportions of each quality, we construct an average-quality index to quantify annual quality volatility over 31 years (1875 to 1912, with seven years missing).

To fix concepts and notation, consider the following price model. At each point in time, observed average market price is P_t (for average quality across all wheat traded at time t). Define P_t^* as the market price if average quality at time t were actually equal to the expected quality (i.e. the quality average across the whole time sample): notice that P_t^* changes over time due to shifts in supply and demand. Then

(1) $$P_t = P_t^* H_t \Rightarrow \ln P_t = \ln P_t^* + \eta_t$$

where $\eta_t \equiv \ln H_t$ represents price effects of quality differences – measurable on American data because we observe prices and quantities of different qualities. For US wheat, standard deviation of η_t (log of quality index) was 0.034 (i.e. 3.4%), with no discernible trend.[2] Putting this into context, measured price volatility – standard deviation of $\Delta \ln P_t$ – of constant-quality US wheat for 1871-96 was 14.6%. Thus annual quality volatility was around one quarter of the magnitude of (constant-quality) price volatility.

Explicit quality data are unavailable for England, or elsewhere, pre-1914. An alternative method uses prices whose variation is unlikely due to anything other than quality variation, such as contemporaneous prices in the same market. Figure 1 illustrates Bristol prices used to calculate the 1790 London Gazette Corn Returns average.[3] Seven weeks saw no trade; seven

[2] For consistency, we use standard deviation of the log of a variable as our measure of volatility (very similar to the coefficient of variation) throughout.

[3] Data from Bristol Record Office, ref 04531(1): Bristol corn inspector's notes and calculations bound into a single volume for 1790 only (other years have only very occasional notes). Bristol wheat price was not published in 1790

saw only one trade; the rest saw more than one, with different trades typically at different prices. For example, in week ending 2 January 1790 there were seven trades at prices ranging from 51/0 per quarter to 61/4 (weighted average of 56/0): about half total trade was a single transaction at 58/4. Since Bristol's cornmarket opened only once per week, these prices are near-contemporaneous and price variation is almost certainly due to quality variation. Estimating quality variation using standard error of weighted mean price, average quality variation is 1.86% of average price. As with US data above, compare this to standard deviation of weekly price changes, at 6.21%: the ratio of standard deviations for these weekly data is 0.30, compared to the US ratio for annual data of 0.25.

Figure 1 here.

How can we measure quality variation, absent explicit quality data or contemporaneous prices? An important source of pre-1914 grain price variation derives from prices being quoted by volume in most countries; many quantitative studies are based on volumetric prices (Keller and Shiue, 2014). Mass is a superior measurement basis because grain mass primarily determines the flour mass produced. Since flour is (and always has been) sold by mass – and since grain value is determined by flour value produced – it makes sense to value grain on the basis of its mass.[4]

Velkar (2012) notes that weighing large amounts of grain was more difficult and costly than establishing volume, pre-1914; hence trading grain by mass was much less popular. The 1834 Returns from Corn Inspectors (BPP 1834, 105) reveal that, of 148 monitored markets, 90 used volume alone; 28 used mass alone; about ten used both; and the remainder are difficult to classify (see appendix A2). However market participants measured grain quantity, data were published in the London Gazette based on volume: a proposal by the Select Committee on the Sale of Corn (BPP 1834, 517) that official returns should include information on both volume

because only county averages were then published. Bristol weekly wheat price standard deviation was 5.9% from the 1820s onwards (when town-level data were published), just slightly lower than 1790's 6.2%.

[4] Flour was often sold by the "sack" in England. Although appearing a volumetric measure, English flour sacks had a standard weight of 240lbs, so were really a mass measure.

sold and density was never implemented. The Select Committee's primary conclusion was that there was significant density variation and this primarily determined grain quality variation. Although most corn inspectors did not provide detailed comments, eighteen inspectors explicitly linked quality to density in the 1834 report.[5] Thus grain density is the main focus of our discussion in this section.

We model cross sectional quality variation using the following decomposition:

$$(3) \qquad \underbrace{P^V}_{\substack{\text{shillings}\\ \text{per bushel}}} = \underbrace{P^*}_{\substack{\text{shillings per}\\ \text{quality-adjusted}\\ \text{lb}}} \times \underbrace{\theta}_{\substack{\text{quality}\\ \text{adjustment}}} \times \underbrace{B}_{\substack{\text{lb per}\\ \text{bushel}}} \Leftrightarrow \ln P^V = \ln P^* + \ln \theta + \ln B$$

where P^V is price by volume; B is grain density; θ is any remaining quality not incorporated into density; and P^* is the price by mass with constant quality. Variations in P^* arise from supply and demand changes.

Federico (2008) proposes variance (or standard deviation) of log-prices as the best market integration measure. From equation (3)

$$(4) \qquad \begin{aligned} \operatorname{var}\left[\ln P^V\right] &= \operatorname{var}\left[\ln P^*\right] + \operatorname{var}\left[\ln \theta\right] + \operatorname{var}\left[\ln B\right] \\ &\quad + 2\left\{\operatorname{cov}\left[\ln P^*, \ln \theta\right] + \operatorname{cov}\left[\ln P^*, \ln B\right] + \operatorname{cov}\left[\theta, \ln B\right]\right\} \end{aligned}$$

Measuring market integration by the arbitrage condition that like goods should have the same price (in markets i and j, $P_i^* = P_j^*$), the Law of One Price should be evaluated using $\operatorname{var}\left[\ln P^*\right]$: in practice nearly all empirical studies are based on $\operatorname{var}\left[\ln P^V\right]$. Note the sigma measure effect in equation (2): parameters may be biased in regression analysis because

[5] Explicit linkage of measurement problems to density and quality are found the Sheffield corn inspector's comment that "the weight per load is often mentioned by the seller in confirmation of the quality of the corn; frequently the small farmers have not the means of ascertaining the weight at home, and then recourse is sometimes had to the scales at the weighhouse in the market." (BPP 1834 (105, p.252), Returns from corn inspectors, 12)

unobserved quality is a form of measurement error and leads to attenuation bias (details in appendix A2).

Many corn inspector returns from 1834 contain bushel weights, summarized in table 1, which offer two quality variation measures. Eighteen inspectors provide bushel weight ranges – typically 6% of average bushel weight: this is how much bushel weights varied within a locality. This number is only indicative, since it is unclear precisely what inspectors meant by this range (variation from different farmers at a point in time, or variation in average bushel weight from year to year?). Forty-six inspectors reported average or typical bushel weight for their area: coefficient of variation for this market cross section is 2.8%.

Table 1 here.

Table 2 shows the relationship between inspectors' reported average bushel weights and London Gazette average bushel prices. Attenuation bias pushes estimated elasticity of price with respect to bushel weight below unity, albeit not statistically significantly so; this is unsurprising, as the bushel weight data are rather low quality and we cannot control for other quality factors.

Table 2 here.

Inclusion of regional or county dummies has little effect on estimated coefficients, although estimates are less precise. Estimated elasticities are consistent with contemporaries' analysis. In 1834, Richard Page gave the example of high-quality wheat (61.25 lbs/bushel) being worth 11% more than low-quality (57.25 lbs/bushel). Two-thirds of the price differential derived simply from differential grain weights. Sellers could overfill sacks of low-quality grain to make them weigh the same as sacks of heavier grain; this would cost 8% of the value of the low-quality bushel. A 3% price difference would remain, arising from the fact that the low-quality grain gave less valuable flour (Page's evidence in British Government, "Select committee," BPP 1834, 348). A decade later, two experimenters (Hillyard, 1840; Barclay, 1845) cultivated different wheat cultivars and measured bushel weights. They then asked local corn factors what price they would fetch. Differences between the most and least expensive samples suggests elasticities of 0.53 and 0.84 respectively (appendix A5). These results all suggest bushel weight correlates with quality and elasticities are consistent in cross sectional data at around two-thirds.

How important is this for price analysis? The unadjusted r-squared for the first regression in table 2 is 0.153, suggesting that $\text{st.dev.}\left[\ln B\right]\big/\text{st.dev.}\left[\ln P^*\right]=0.43$ (appendix A4 has more details and calculations). So quality variation may have been slightly greater in 1830s England than in the US at the end of the century (which we estimated above as about one quarter).

Density is the primary – but not sole – quality indicator. Wheat type (red/white, hard/soft, winter/spring) was an important determinant and affected price. Jago (1886, 233) notes spring wheat contains more starch than winter wheat. “Hard” wheat contains more gluten, necessary for good bread but not biscuit. This raises two questions. Would density be a sufficient statistic to enable dealers to judge quality? To what extent is density important for quality *per se*, and to what extent correlated with other quality characteristics?

Evidence that density was a sufficient statistic comes from the 1834 Select Committee (BPP 1834, 517). Layton Cooke, Chairman of the Committee of Agriculture in the London Society of Arts, stated:

> “A considerable dealer in corn called upon me this morning ... : I stated to him that there was an idea of adding weight to measure [volume]; he said, then if that is the case, then skill would be of no value; anyone might buy corn as well as ourselves.”

Were other quality characteristics correlated with density? Jago (1896) provides an 1881 cross section of English and imported wheats. In table 3, we regress price per bushel on bushel weight alone, and then on bushel weight and other quality determinants (full results in appendix A5). Variables are logged, so parameter estimates are elasticities and comparable to table 2.

Table 3 here.

Bushel weight was evidently correlated with other quality determinants, since additional explanators (gluten, impurities) reduce the estimated density coefficient. Elasticity of price to density is much higher in table 3 than our previous estimates, presumably because wheats from different countries varied more than wheats within England.

Eighteenth century English commentators talked about grain characteristics (density, cleanliness, soundness of kernels) and their price effects. A fundamental part of grain trading was sampling grain parcels for these characteristics to determine market price. High impurities, for example, made grain unsalable even in the eighteenth century (Ellis, 1744). Nathaniel Palmer (1834) speaks of wheat being “too hot” (starting to ferment) from being stored damp, and then being brought "into condition" by turning and aerating. In 1814-33, the volume lost by 690,000 quarters of stored wheat coming into condition averaged 2% (appendix A5). Absent systematic evidence on these grain quality dimensions, we can say only that wet harvests would result in more grain that was "too hot" being brought into granaries.[6] It is impossible to trace quality and condition across space and time.

So return to the primary quality determinant – bushel weight – about which we can say something well founded and rigorous. Table 4 reports "average" or "representative" bushel weights for wheat, barley and oats (additional grains in appendix A4.C). These are generally based on an official source – such as the Corn Returns Act, or the trade and navigation accounts. There is an upward trend, especially marked for wheat and barley; bushel weight increased by perhaps 7% between 1791 and 1902.

Table 4 here.

Now consider the variance around these averages. Table 5 provides evidence on both within-harvest and between-harvest variation. Wheat bushels in an average year weighed around 59 pounds (consistent with the Corn Returns Act), good years being around two pounds heavier and bad years around two pounds lighter. Within-year quality variation was larger, with high or low quality wheats weighing perhaps five pounds more or less than average.

Table 5 here.

We charted long run bushel weights changes, year-to-year differences and within-year variation. How much flour – and what type – did each pound of wheat produce? Flour content ultimately defines grain quality. Variations in flour quantity, or quality, may explain grain quality variation not due to bushel weight. Although varying slightly in the details, English flour was generally assigned to one of five or six categories. "Household" was best quality, used for baking white bread commonly eaten in London. "Seconds" was mixed with Household, or used by bakers who sold bread below the maximum price. "Thirds" were

[6] Henry and Kettlewell (1996, 430-1) offer a modern analysis. Jago (236-37) notes that damp wheat was disadvantageous for two reasons: it effectively meant purchasing water; it was subject to mustiness.

shipped out of London, used for brown bread in the provinces. “Fourths” (sometimes divided into “Fine middlings” and “Coarse middlings”) went to Liverpool or Newcastle for ships’ biscuit. Pollards and bran were not used for human consumption (Bennett and Elton, 1898). Component proportions varied significantly across parcels of wheat, greatly affecting value. Thomas Dimsdale, corn factor, was asked by a Parliamentary Committee: “What proportion do you reckon that a sack of flour compares to a quarter of wheat, generally speaking?” He answered:[7]

> “It depends so entirely upon the quality of the wheat, that I should mislead your Lordships by giving an answer. Flour, if good, will make more loaves per sack than if indifferent.”

Table 6 quantifies quality variation, reporting how much flour (of each type) derived from bushels of different weights.

Table 6 here.

The 1841 data are based on flour output of 13 grain samples. Although small, the sample permits regression analysis (appendix A5). Regressing log of flour extraction rate on log of bushel weight reveals 1% increases in bushel weight raise the proportion of flour by 1.4% (t-statistic=2.29, r-squared=0.26).

Table 3 extraction rates are high compared to other sources. Sir T. H. Elliott’s 1903 Cabinet Memorandum, quoted by El-Husseini (2002), suggests flour extraction rates of 72%; Petersen (1995) suggests 70-75%; Feinstein used 75% (personal communication). Although our rates are higher than other authors’, the sources we found are consistent with one another and generate a plausible pattern. Panels B and C show heavier bushels contained higher proportions of farinaceous material. Heavier bushels also have larger mass (by definition), so two effects pushed up high quality bushels’ flour content – larger mass, and higher farinaceous proportion. This correlation is useful because bushel weight captures not only variation in grain mass, but also variation in flour mass – the fundamental determinant of bushel value – and makes bushel

[7] British Parliamentary Papers (1826-7), “Report”, 674.

weight closer to being a sufficient statistic for quality variation. The outlier is Panel A, having light bushels but high extraction rates. Those data are drawn from London's Albion Mill, England's first steam-powered mill; perhaps the powerful, new machinery was better able to separate farinaceous material from bran, pollard and waste – hence having higher extraction rates. It is also consistent with figures from the Parliamentary enquiries around that time: Samuel Kingsford, miller, gives typical flour extractions rate of 81.25% (British Government, "Report", BPP 1813-14, 292). Fire destroyed the Albion Mill in 1791, but the nineteenth century switch to steam milling possibly pushed up flour extraction rates.

Consider how grain quality affected price. Table 7 reports 1834 evidence from Stead. Prices *per pound of grain* are similar for grains of all qualities: first quality sells for around 5% more per pound than third quality, in both good and bad years (consistent with heavier bushels having proportionately more flour). Obviously, bushels weighing 10% more would sell for 10% more: more grain mass generates more flour mass. But bushels weighing 10% more actually sold for 15% more, owing to higher proportions of farinaceous material. Hence *price per pound of grain* was 5% higher for heavier bushels. This evidence suggests that probably two-thirds of grain quality variation arose from density variation.

Table 7 here.

3. Cross-sectional quality variation

Many factors generate grain quality variation across England. For example, storage conditions affect grain weight (through moisture content) and condition. But cultivar was probably the most important systematic grain quality variation determinant. Cultivar could vary due to supply or demand. Consider supply. Some cultivars suit certain climate and soil conditions better than others: thus cultivars grown in the drier east differ from those in the wetter west, for example. Also, there was a possible trade-off between grain and straw production: farmers located further from grain markets, producing relatively more animals, had relatively higher values for straw and might rationally choose lower-yielding cultivars. Finally, we assume nowadays innovations spread rapidly: less obviously true in the eighteenth century, superior cultivars might take years to diffuse. Parliamentary enquiries reveal significant systematic wheat supply quality variation. John Coupland, corn factor, was asked: "As a corn factor, if you knew that the price of wheat in the Lincoln market was 60 shillings, would you, in giving an order for foreign corn, calculate upon an importing price of 60 shillings or above it?" He answered (British Parliamentary Papers 1826-7, "Report", 738):

> "I certainly should not expect to receive 60 shillings for what I imported. I should conceive that the average of the Lincoln market is much above the average of the country. The wheats in the Lincoln market are much better than the average quality; they are better by several shillings than the average quality."

Now consider demand. Different places value seed characteristics differently: high-yielding cultivars are valued in potential food shortage areas, but high quality cultivars where people consume fancy baked goods (London). John Hodgson, Liverpool corn merchant, was asked: "Were those [two] wheats of nearly the same quality?" And he replied (British Parliamentary Papers 1826-7, "Report", 753):

> "There is considerable difference in the quality; the quality of wheat imported into Liverpool is generally much inferior to the qualities imported into London; the wheats in question were 6 shillings to 8 shillings inferior to the quality of that which is sent to London; the consumption of Liverpool being generally of an inferior description of wheat; to place Liverpool and London qualities on a par, fully 6 shillings [20%] must be added to the Liverpool prices."

This impacts the market integration literature because systematic quality variation prevents prices equalizing across markets.

Remarkably few historians have analyzed cultivar effects in English agriculture. Walton (1999) offers a largely descriptive treatment of changes over time. Anecdotal evidence reveals different cultivars being grown across regions, but systematic evidence is unavailable. An author may state Rivet is popularly grown in Lincolnshire, being better suited to the climate and soil. But we do not know *all* farmers there were growing it; or whether cultivation extended *throughout* Lincolnshire; or whether its use was *exclusive* to Lincolnshire. On the contrary, farmers often sowed several cultivars to suit different conditions around the farm (Ellis, 1744, 33-4; Trowell, 1750, 9); this also provided weather insurance and staggered harvest dates, spreading labour demand at peak time. Such heterogeneity makes it difficult to quantify cultivar effects on yield differentials between Lincolnshire and elsewhere. However, presenting the available data puts bounds on the problem: how much *could* yields have varied around England from cultivar differences?

Rothamsted experimental farm undertook cultivar experiments, 1871-81, to quantify yield differentials and test whether newer cultivars gave higher yields than traditional ones. This agenda is key because must consider how traditional cultivars impacted regional yields and quality in the eighteenth and early nineteenth centuries. Absent contemporary experiments, our best approach is to examine later experiments based on the same seed stock. This measure is imperfect because quality of a given cultivar may have improved. Eighteenth century farmers employed "in-breeding" – taking the best kernels from their current crop, and sowing only those, to propagate strains with desirable attributes. In 1601, Maxey lauded high yields from "well-dunged land sown with choicely picked seed". In 1788, Marshall recommended using the best ears as special seed stock. The Romans used "mass selection" of the best ears, whilst "pedigree selection" (in-breeding) was used in the nineteenth century; Percival (1934, 43, 75, 83-4) discusses these points. Possible quality improvement makes it hazardous to take nineteenth century cultivar data and project it back to the eighteenth century. However, we pursue this approach because no systematic eighteenth century data exist (i.e. we have little alternative) and we are careful to adduce corroborating qualitative evidence.

The *Journal of the Royal Agricultural Society of England (JRASE)* contains cultivar data from around 1841. The Society's motto – *Science with practice* – was implemented from the outset by organization of cultivar trials across England. Members sought to establish the best cultivars, and thereafter raise average yields by popularizing them. The Society's trials were less systematic than Rothamsted trials, taking place on private, working farms. This reveals how cultivars performed in a range of realistic farming conditions (unlike Rothamsted trials); but trial heterogeneity makes it difficult to assess whether results were driven by farming practices or cultivar characteristics. Different farms had different soils, climate, manures, crop rotations and sowing practices (drilling, dibbling or broadcasting; sowing thickly or thinly; and early or late in the season). Large samples might balance out such variations but most trials involved few cultivars and we typically have only a few trial observations for each cultivar. We analyze data from the seven largest trials (each testing 5-17 cultivars).

Eighteenth century information derives from Ellis – a well-informed practical farmer, agricultural writer and seedsman who gives considerable detail on numerous cultivars of wheat and other crops. The evidence is not quantitative but permits us to trace the history of different cultivars and get qualitative descriptions – such as Red Lammas being considered superior for bread making in 1750, just as in 1850. Appendix A4.F reports yields and bushel weights for all identifiable cultivars in 1750, 1841, 1871 and 1914. Here we summarize the most striking features.

Rivet, said to be the oldest English wheat still cultivated in the late nineteenth century, gave heavy yields on strong land but had coarse straw and low quality grain. Walton (1999, 49) notes Rivet's exceptionally low gluten content, making it inferior for bread. Rothamsted results confirm Rivet's high yields, giving 21% more bushels/acre than Nursery Red, one of the lowest yielding cultivars. But Rivet's quality was indeed low, bushels weighing 7% less. So – controlling for bushel weight – Rivet was only 14% more productive. Overall, yield and quality are strongly and significantly *inversely* correlated across the Rothamsted sample (-0.6, p=0.01). Farmers could choose high yields and low quality, or vice versa, but not have both; contemporaries noted this (Percival, 1934). It may explain persistence of so many cultivars: one cultivar was not superior, it simply offered a different quantity-quality trade-off. Pooling the two largest *JRASE* trials, the yield-bushel weight correlation is -0.4 (p=0.08, N=21). These two trials were carried out in different years – one in wet Gloucestershire, the other in dry Lincolnshire – so we find it remarkable that results are so similar to Rothamsted.

Now consider yield levels. Average yield across all cultivars was 42.5 bushels/acre, 1871-8. The average for the three known eighteenth century cultivars (Rivet, Red Lammas, Golden Drop) was 46.7 bushels/acre. New seed cultivars generated no obvious yield increase, 1750-1871. We cannot be certain because post-1750 disappearances – White Cone, Red Pirkey and so on – perhaps gave lower yields, so the overall average could be raised by discontinuing unproductive cultivars. We can only be sure the upper tail of the distribution was unchanged.

Table 8 below compares Rothamsted and *JRASE* data, reporting all cultivars that we can match. Yields are predictably different (Rothamsted averaging 17% higher, and a correlation between the two yield samples of 0.5, p=0.32). But bushel weights are almost identical (Rothamsted averaging 0.2% higher, and a correlation between the two sets of bushel weights of 0.7, p=0.15). Take two cultivar data sets, 30 years and hundreds of miles apart. You cannot easily predict yield levels – but you can predict yield rankings well, and bushel weights with extraordinary precision. Bushel weights in 1914 were also similar to 1871. This characteristic is very robust.

Table 8 here.

Such persistent quality variation affects attempts to quantify market integration. The market integration literature presumes Law of One Price (LOOP) holds when transport costs are zero. But LOOP will *never* hold: if Kent-Lincolnshire transport costs were zero, wheat would not trade at the same price because quality differed. How large was English cross sectional quality variation? Take average wheat prices for 1886-1914 for each county – long enough to smooth out random fluctuations, and so truly capturing equilibrium county prices. The coefficient of price variation is 2%. The coefficient of quality variation, *c*.1871, is also 2% (appendix A4). So quality variation *could* explain observed price variation. We matched seven cultivars to particular counties, based on qualitative literature: Cumberland (Fenton), Essex (Essex Brown), Gloucestershire (Bristol Red), Middlesex (London Red), Northumberland (Hopetown), Somersetshire (Bristol Red), Surrey (Surrey White). The quality-price correlation is 0.6 (p=0.16). Again, the evidence is weak – owing to few observations – but is consistent with quality determining long run price variation across English counties.

International comparisons are particularly problematic. Persson (2004) and Hynes *et al.* (2012) tackle this by restricting their analysis to a particular product (such as "Manitoba No. 2 Red Wheat"); but this is feasible only for the late nineteenth century Atlantic trade, not for earlier periods or less developed markets, where products were less tightly defined. Estimates of price dispersion (favoured in Federico, 2008) are especially problematic. We have considered quality variation across England only – a small and relatively homogenous locality. The quality problem increases for larger regions (India, China) and international markets; it may also be larger for other commodities (perhaps rice?); and geographical quality trends can generate price convergence unrelated to market integration, as in the next section.

4. Long-run quality variation

English and foreign wheats differed in quality. Table 9 reports international imported bushel weights; the dramatic shift in English wheat import provenance necessitates changing geographical focus between 1825 (panel A) and 1900 (panel B). Note the widespread bushel weight increase, 1825-1900, and marked convergence (i.e. wheat became more homogenized).

Table 9 here.

How did bushel weight affect market values? Johnson (1902, 2) states explicitly grain parcels are entered into trader's ledgers pro-rated. A trader accepts 100 bushels weighing 62lbs/bu, rather than the normal 63lbs/bu. This enters the ledger as a stock of 98.4 bushels (=100×62/63), while the bushel price is reflated by 63/62. This implies the elasticity of price to bushel weight was 1 (bushels 10% lighter were worth 10% less and *vice versa*). Weeks (1871) suggests this was standard accounting practice by 1871, at the latest. This supports evidence above regarding high and low quality wheat within England (a one-to-one relationship between bushel weight and market value).

Consider earlier years. Using total values and grain import quantities, we can calculate weighted average prices of foreign grains traded in England. All these prices are net of freight charges and import duties (i.e. these are the prices you pay at Mark Lane), so price variation can reflect only quality variation. Data were reported in bushels to 1863 – useful because English Corn Returns data are on the same basis and directly comparable – and reproduced in table 10, column 2. North American wheat quality equalled English; German was slightly better; western European was slightly worse; southeastern Europe and eastern Mediterranean – especially Egypt – were significantly lower quality.

Table 10 here.

Post-1863 trade accounts claim to report quantities in hundredweights, which is almost certainly false. Grain was traded in England in bushels, not hundredweight. So conversion of imported grain from bushels to hundredweights requires data on lbs/bu for each parcel – information unavailable to the Government. Also, most quality variation arises from bushel weight variation: reporting prices per hundredweight, rather than per bushel, would eliminate most price variation. Post-1863 variation is actually the same as 1855-63 variation (compare table 10, columns 2 and 4); correlation between the two cross sections of country average prices is 0.89 ($p<0.01$). How did the Government generate the hundredweight data? We are not told but can infer it. An 1887 account reports imports back to 1866 in quarters, and notes they have been converted from hundredweights at 4.5 hundredweight/quarter (63lbs/bushel) (British Government, “Wheat and flour imports”, BPP 1887, 300). Taking hundredweight data for each country, and converting it back into bushels at 63lbs/bushel, we get the prices reported in table 10, column 4. Prices are slightly lower in 1864-74 than 1855-63, but differentials are extremely similar. Egypt is exceptional, as its quality converged to the Eastern Mediterranean average.

This conversion saga has several implications. First, suppose the Government were correct that imports averaged 63lbs/bu. Total import estimates, expressed in hundredweight from 1864 onwards, would be correct. But individual country figures would be wrong: accurate country figures require country-specific bushel weights. Second, reversing the Government's calculation enables inference of number of bushels imported from each country, and then prices per bushel (since we know total import value from each country). We can thereby chart quality changes over time. Third, weighting the prices by trade volumes enables us to calculate changes in the average consumption price of wheat in England.

Figure 2 charts international wheat quality changes by plotting price ratios of European wheats to English wheat. Ratios exceeding one imply foreign wheats were higher quality than English wheat.

Figure 2 here.

All series are flat (trendless compared to English wheat). Danish and French wheats were consistently slightly lower quality; German and Spanish wheats were consistently slightly higher. Danzig furnished much of the German wheat sent to London, and was renowned for high quality (Claude Scott, corn factor, in British Government, "Minutes", BPP 1795, 26; John Lander, Maltese Government Agent, in British Government, "Report", BPP 1826-7, 649, and John Birkett, corn factor, 657-8, 660; Capper, *Port*, 230-1). Double-checking the half-dozen very low values reveals no errors, but those prices typically pertain to small quantities; they could be small loads of very low quality grain (perhaps damaged in transit), or maybe data recording errors at the customs house. There is some suggestion that English wheat quality fell after 1885 (Biffen, "Mendel's laws", 4-5; Percival, *Wheat*, 70-1). If so, quality of other European wheats was falling at a similar rate. Indeed, this is plausible because western European nations bought English seed (Humphries and Biffen, 1907-8, 2-3).

Figure 3 charts New World wheat quality. There is clear upward movement, from a price relative of 1 to 1.1; a discrete step around 1878 moves all the series up together. Australian wheat was significantly higher quality than others; Indian wheat started out lower but converged by 1914.

Figure 3 here.

Figure 4 is most dynamic, with a strong upward trend from 0.8 to 1.1, 1855-1914. Egyptian wheat started lowest but gained most; Persian wheat was highest. Eastern Mediterranean wheat was long known in England for low quality (John Wilson, corn factor, in British Government, "Report", BPP 1813-14, 273; John Lander, Maltese Government Agent, in British Government, "Report", BPP 1826-7, 652). A key problem was poor threshing technique (livestock treading); it added impurities to grain, which was not then cleaned and went out of condition faster (S. Bosanquet, Governor of the Turkey Company, in British Government, "Minutes", BPP 1795, 26; George Baldwin of Alexandria, in British Government, "Minutes", BPP 1795, 183; Thomas Dimsdale, corn factor, in British Government, "Report", BPP 1826-7, 666, 676; and John Schneider, Russia merchant, in British Government, "Report", BPP 1826-7, 730.) There was likely improvement in processing in the late nineteenth century eastern Mediterranean, which increased quality and market value in London. Agriculture expansion to virgin soils in the Great Plains likely raised average American grain quality (in both North America and Argentina). By contrast, western and central European soils – cultivated for 1 000 years and already well-managed in 1855 –saw no quality improvement to 1914.

Figure 4 here.

Our second exercise is constructing the weighted average price of grain actually consumed in England (combining foreign and domestic product); this is the "consumption price" of wheat. The trade accounts, following the hundredweight/bushel conversions discussed above, enable this. Figure 5 charts consumption price, together with weighted average English price from the Corn Returns. English prices tracked import prices closely, as expected in a well-functioning market. It seems there would be no substantive difference using consumption or English prices for cost of living indices. Such obvious inference is sadly incorrect. Consider price differentials, instead of levels. The English price premium (right scale of figure 5) declines markedly, 1839-1914. Foreign prices rising faster, and their increasing market share, means the consumption price rose significantly faster than the English price – 0.1% per annum (t-statistic=7.11, r-squared=41%). This is a lot in a consumer price index over 65 years. Thus it is important to use price series close to the item of interest when constructing the index; Feinstein's index (using bread prices) is likely more accurate than Clark's index (using English wheat prices).

Figure 5 here.

We documented substantial quality differentials between English and foreign wheats (New World wheats were better, eastern Mediterranean wheats worse). We charted significant quality changes over time (many wheats improved quality, relative to England, especially Egyptian and eastern Mediterranean). The falling standard deviation of log prices in table 10 (0.117 to 0.092) may look like improved market integration, but is due entirely to quality convergence.

Also, rising foreign prices and market shares means the consumption price of wheat (per bushel) rose significantly faster than implied by the Corn Returns.

5. Measurement error effects in time series models

If quality variation generates invalid inference of market integration using coefficient of price variation, are we safer using time series methods commonly used? No. Most modern time-series studies use a cointegrating VAR framework of the form

(5) $$\Delta \mathbf{p}_t^* = \boldsymbol{\alpha}\boldsymbol{\gamma}\mathbf{p}_{t-1}^* + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \quad \forall t \neq s : \mathrm{cov}\left[\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_s\right] = 0, .$$

where $\mathbf{p}_t^*$ is a vector of log-prices for constant-quality wheat, as defined in equation (1). With cointegrated prices, market efficiency (speed of return to equilibrium) is measured by estimates of $\boldsymbol{\alpha}$ (the "loadings"). These can be estimated alongside estimates of $\boldsymbol{\gamma}$ (Johansen procedure) or by imposing price homogeneity (OLS).[8]

Estimates of long-run relationships are unaffected by measurement error. But here we quantify measurement error effects on causality direction and estimated market efficiency (speed of adjustment) when observed volumetric prices $\mathbf{p}_t^V$ are affected by measurement error of form

(6) $$\mathbf{p}_t^V \equiv \mathbf{p}_t^* + \boldsymbol{\eta}_t .$$

For simplicity, consider the case with two prices, A and B, and price homogeneity so that $\boldsymbol{\gamma} = \begin{pmatrix} 1 & -1 \end{pmatrix}$, and remaining assumptions about disturbances and errors are:

[8] Two further issues are not discussed here. One can test for market integration (price cointegration) using the Johansen trace test; one can estimate the cointegrating relationship summarized by vector $\boldsymbol{\gamma}$. Hassler and Kuzin (2009) show the cointegrating vector can be estimated consistently and the Johansen test continues to be reliable (probability of making a type I error is unaffected) provided sufficiently many lagged price differences are included in the VAR to remove serial correlation in residuals. Second, Nielsen (forthcoming) introduces a technique to measure the speed of adjustment when there is measurement error.

(7) $$\begin{aligned}&\operatorname{corr}\left[\varepsilon_t^A,\varepsilon_t^B\right]=0.5;\quad \operatorname{var}\left[\varepsilon_t^A\right]=\operatorname{var}\left[\varepsilon_t^B\right]=\sigma_\varepsilon^2\\&\mathrm{E}\left[\boldsymbol{\eta}_t\right]=\mathbf{0};\quad \operatorname{var}\left[\boldsymbol{\eta}_t\right]=\sigma_\eta^2\mathbf{I};\quad \operatorname{cov}\left[\boldsymbol{\eta}_t,\boldsymbol{\eta}_{t-1}\right]=\mathbf{0}\end{aligned};$$

There is long run price homogeneity, some correlation between disturbances to the true price series (invariably observed in real data), and classical measurement error in both price series. Disturbances and measurement errors are assumed Normally distributed. We considered a variety of parameter values for equation (5) and the measurement error, but report results here based on:

(8) $$\boldsymbol{\alpha}\in\left\{\begin{pmatrix}0\\\alpha\end{pmatrix},\begin{pmatrix}-\alpha/2\\\alpha/2\end{pmatrix}\right\};\quad \sigma_\eta\in\left\{0,\sigma_\varepsilon/4,\sigma_\varepsilon/2,\sigma_\varepsilon\right\}$$

Consider two possible loadings configurations. First, "asymmetric loadings": price A Granger-causes price B (only price B adjusts to remove disequilibrium); so Uxbridge adjusts to London, for example. Second, "symmetric loadings": both prices adjust the same amount to remove disequilibrium; so Alnwick and Berwick adjust towards each other. We choose values of α to determine shock half life. We analyze four possibilities for measurement error magnitude (whose standard deviation is compared to standard deviation of disturbances to the underlying price). Earlier, we compared measurement error standard deviation to that of observed prices, S_h/S_p: we report this relationship in table 11 when half life is two. This is lower than the ratio $\sigma_\eta/\sigma_\varepsilon$ because the measurement error induces a negative moving average error in equation (7). Our results above suggest relevant simulations are when $\sigma_\eta/\sigma_\varepsilon$ is between a quarter and a half.

Table 11 here.

Figures 6-8 summarize measurement error effects on estimation by plotting median impulse response functions when half life is two time periods and sample size is 200. Top-left curves, calculated with no measurement error, are very close to the true impulse response functions (slightly wrong because a sample size of 200 is not quite enough to remove all small-sample bias). Figures 6-7 show measurement error effects with asymmetric loadings (Uxbridge-London). Shocks to A should be permanent, but some of the effect dissipates with measurement error; and, although shocks to B should completely disappear, a small part of the shock appears permanent. Bias is largest for shocks to the exogenous price. Figure 8 shows the effect for symmetric loadings (no need to produce two sets of graphs here, as loadings symmetry implies identical impulse response functions). The bias here is small, even with large measurement errors.

These simulations show measurement error must be larger than we observe in our data to generate large biases in estimating market responses – except if one market dominates another, when estimation results erroneously suggest bi-directional causality.

Figure 6 here.

Figure 7 here.

Figure 8 here.

But convergence speeds are greatly over-estimated when unobserved quality variation generates price variation measurement error. In figures 2-4 above, compare top-left graphs (no measurement error) and bottom-left (measurement error as large as we observe in English data). Estimated shock half life falls erroneously from two weeks to one.

Table 12 reports measurement error effects on processes with different half lives, comparing true and estimated values. Half lives are always biased down (even with no measurement error) but bias increases both in the measurement error and the true half life. The reason is that apparent convergence to equilibrium is very fast when measurement error is significant: when $\sigma_\eta / \sigma_\varepsilon = 1$, typically less than half the disequilibrium remains after only one period, regardless of true half life.

Table 12 here.

So the longer the estimated half life, the larger the bias. Many analyses suggest half lives of several months, even half a year (Persson, 2000). Our Monte Carlo simulations suggest the true half life is only half as long.

6. Year-on-year quality variation

Inter-year grain quality variation is potentially important. We demonstrate positive correlation between yield/acre (quantity) and quality. So true grain quantity – tonnes of grain or flour – was greater than the bushel measure in good years, and lower in bad years. How does this bias, for example, estimated elasticity of demand?

Rothamsted furnishes two data sets to help us. Cultivar trials, 1871-81, provide yields and bushel weights (graphed in appendix A5.A); so do continuous wheat experiments started in 1844 (Lawes and Gilbert, 1864; 1884) (graphed in appendix A5.B). The latter are extremely detailed, reporting dressed corn average bushel weight, as well as total dressed corn weight and total offal weight. We analyze two time series, 1844-83: the plot manured with yard dung, and the plot remaining unmanured. We chose these plots as best reflecting actual conditions on working farms, where – to 1860, at least – most land received yard dung or nothing. Table 13 reports results for the manured plot and pooled regressions for both plots (the unmanured plot is presented separately in appendix A5). Since neither slope coefficients nor intercepts differ significantly between manured and unmanured plots, the pooled regression is our best estimate. We estimate all regressions for the whole period and for a sub-sample omitting 1853, 1879 and 1880 (years with particular problems and low yields).

Panel A reveals positive and statistically significant time series yield-bushel weight correlation. Omitting exceptionally bad outlier years results in lower estimated elasticities of about 4% in typical years. Panel B reports results for different cultivars. Walton (1999) suggested a negative yield-bushel weight correlation across cultivars; we find a large negative elasticity (-0.11) using the largest possible balanced panel (1873-78), although it is statistically insignificant. Overall, the relationship is positive. This need not suggest some cultivars dominated others: some may be better adapted to alternative soil types, or produce more straw. The within-group estimator reveals similar results to the time series analysis: there is a positive elasticity of 5%, rising to 11% if exceptionally bad years are included.

Table 13 here.

Suppose elasticity of density to yield was 10%. What does this imply about bushel weight fluctuations? National agricultural returns report average English wheat yields of 30.7 bushels, 1885-1914, with a standard deviation of 2.6 (8.5%). Thus a harvest two standard deviations above average would see yields 17% above average, implying bushel weight increase of 1.7% (=0.17×0.10), or 1lb if bushels averaged 60lbs. This is somewhat below eighteenth century estimates, reported in tables 4 and 6, when commentators suggested bushels weighed 2lbs above average in good harvests and 2lbs less in bad. Of course, eighteenth century yield volatility may have been higher. Widespread mid-nineteenth century clay pipe drainage installation may have reduced volatility (Phillips, 1989); pipe drainage removed excess water more effectively and led to smaller crop losses in wet years, lowering yield volatility after the

1830s. In any case, our 10% estimated elasticity of density (quality) to yield (quantity) is more likely too low than too high.

Now estimate wheat demand elasticity *adjusting for year-on-year quality variation*. We find elasticity estimates 10-15% higher than estimates based on unadjusted quality. Barquín (2005) provides an exhaustive study of European wheat demand elasticities from the late mediaeval period to 1914. He incorporates all previous estimates, such as Fogel's and Persson's, and suggests a late nineteenth century English elasticity around 0.43, or 0.68-0.78 for other countries (Barquín 2005, 260-1). So year-on-year quality variation adjustment would push estimates to 0.5 for England and 0.8-0.9 for other countries. This is about the same magnitude as adjustments for carryover and seeding rate biases, which Barquín makes. But those two biases offset each other, whereas quality variation pushes estimated elasticities decisively downwards.

7. Intra-year quality variation

Several studies analyze intra-year grain price patterns (McCloskey and Nash, 1987; Clark, 1999; Brunt and Cannon, 1999). Theoretically, grain price paths should be saw-toothed: starting from a post-harvest minimum, rising gradually through the year, dropping abruptly at the next harvest. Why? Those holding grain through the year must be compensated for it – receiving appreciation on their stocks to offset storage cost (granary rental), storage losses (to vermin) and opportunity cost (return from investing the capital elsewhere). This conceptualization has been used to justify inference of local rates of return on capital – steeper grain price rises through the year imply higher local interest rates. What if we incorporate quality variation into the analysis?

The best grain was never marketed but retained – or sold privately – for seed (Ellis, 1744, 339-40); Trowell, 1750, 9; John Porter, Office of the Committee of Privy Council for Trade, in British Government, "Minutes", BPP 1795, 186). Winter (1798, 131) states seed grain commanded a 10% premium. England never suffered a famine requiring farmers to sell their seed corn for milling, so we never see it in the official markets. The worst grain (offal, or tail corn) comprised smaller kernels, perhaps broken and contaminated with non-wheat seeds (Ellis, 1745, 129). Offal was typically consumed on-farm. But if the harvest turned out lower (or demand higher) than farmers expected then offal coming onto market later in the year could put downward pressure on (rising) prices because those bushels would be lower quality. Then we would systematically underestimate price increases in high-priced years.

Brunt and Cannon (2013, 324) suggest offal constituted 6.5% of the harvest, and 10% of output was retained on-farm in 1801 to feed farm families. Thus we expect zero offal to come to market. But 1851 on-farm consumption amounted to perhaps only 5.3% of total consumption, given changes in agricultural population (British Government, *Census*, vol. 1). By 1845 – the eve of free trade, when grain imports were perhaps 10% of consumption – 5.3% of consumption amounted to around 5.8% of domestic output. Thus on-farm consumption probably still absorbed all offal, especially since some was used for fattening livestock (Walton, 1999).

What effect would offal have on market prices? Barclay's (1845, 192-3) price quotations for dressed grain and offal, for five different wheat cultivars, reveals offal selling for an average 14.7% discount. If only offal were marketed in the final month before the harvest (an extreme assumption) then prices would drop 14.7% – approximately equal to the average intra-year price increase (Brunt and Cannon, 2012). Thus the unwary might estimate zero annual rates of return on grain holding, whereas it was truly 15%. This is not problematic with English data because significant offal probably never came to market. Also, Brunt and Cannon avoid price data from late in the harvest year, when they are most likely to be contaminated. But bias could arise in other countries and circumstances, such as famines.

Intra-year quality variation certainly impacts barley. Brewers bought best quality malting barley soon after harvest, when the market was most active (Brunt and Cannon, 2013). Trading fell sharply later and most activity was in lower quality (non-malting) barley; price impact was large. Price courants report malting barley (several cultivars) and other barley. For example, Tuesday editions of the *Courier and Evening Gazette* list prices of many types of grain and seed, including barley, fine barley, malting barley and fine malting barley; in 1799 they traded around 38, 41, 45.5 and 49 d/bu respectively. The *Morning Chronicle's* "Corn Exchange Report" in 1841 has grinding and malting barley at around 22.5 and 29 sh/qu respectively. With malting barley commanding a 20% premium, trading all malting barley in the autumn – and all grinding barley in spring – generates a 20% price decline. This is offset by intra-year price appreciation acting as a holding return. But without good data on the malting mark-up and the monthly share of malting barley in total trade, we cannot separate these two effects. Thus barley prices are useless for inferring holding returns.

Oats may mirror barley. Horses consumed oats; so did humans (particularly in northern England and Scotland) as porridge and oat bread. High quality oats went for human consumption (Nathan Palmer in British Government, “Select committee”, BPP 1834, 258); so, perversely, high quality parts of the harvest traded systematically in low-income areas (whereas one might expect it would go to richer areas, like London). It is likely there was systematic oat quality variation through the year – human supplies secured first (at high prices) and horse oats later (at lower prices). It seems sensible to treat intra-year oat price movements with some caution.

8. Conclusion.

We examined wheat quality variation across England and over time (intra-year, inter-year and long run), and international quality differentials. Quality differentials arise primarily from bushel weight differentials – useful because they are quantifiable. Contemporaries could assess quality variation by inspection, so market prices captured them.

Inter-year quality variation was relatively small. Quality and yields being *positively* correlated annually, variation in quality-adjusted wheat output was around 15% higher than unadjusted variation. One can control for this in time series analyses – such as estimating demand elasticity – given available quality data. This would increase estimated demand elasticities by 15%, compared to previous estimates.

There was marked cross sectional quality variation *inversely* correlated with local yields (places generating high volumes produced low quality). The long run stability of this pattern in England – the same counties grew the same cultivars for centuries – suggests it was an equilibrium. Some localities optimally chose high quality (near London, where quality fetched a premium); other localities optimally chose high volume (Lincolnshire, where they made ship’s biscuit). Cross sectional quality variation implies the Law of One Price would never hold strictly – prices never fully converging even with zero transport costs. Market integration measurement using coefficients of price variation are problematic; the wider the net is drawn (international versus national versus local), the more quality variation we will see, and the further we are pushed from the Law of One Price (irrespective of transport costs). Moreover, international data show the size of this effect changed over time – marked international quality differentials in 1825 and 1855 vanished by 1914 – which generates spurious evidence of market integration.

Unfortunately, transient random quality shocks through the year bias market integration measurement using error correction models because price responses to quality shocks are confounded with responses to price shocks. Wheat quality volatility was sufficiently high that half lives have likely been overestimated by 100% in the literature.

Systematic intra-year quality variation was likely not problematic for English wheat. But barley quality declined markedly through the year, pushing prices down 20% between one harvest and the next. Oat quality likely declined through the year, too. So English wheat prices offer a safe basis for inferring rates of return on grain, but barley and oat prices do not. In other countries or time periods, systematic variation in intra-year wheat quality cannot be discounted *a priori* and must be considered when analyzing price variation.

Further research quantifying grain quality variation over space and time (using price courants, the agricultural census, government price data) would greatly increase precision and reliability of estimates of demand elasticity and market integration.

Primary sources.

Ardlie, Robert, Evidence recorded in appendix 6 of British Government, "Fourth report on the grain trade," BPP 1797-8. See Lambert, Sheila (ed.) (1975), *House of Commons sessional papers of the eighteenth century* (Delaware: Scholarly Resources Inc.).

Ball, C. R., C. E. Leighty, O. C. Stine, and O. E. Baker (1921) *Wheat production and marketing* (Washington: Government Printing Office).

Barclay, David (1845) "On the advantage of thick sowing," *Journal of the Royal Agricultural Society of England*, vol. 6, 192-3.

Biffen, R. H. (1905-6) "Mendels laws of inheritance and wheat breeding," *Journal of Agricultural Science*, vol. 1(1), 4-48.

British Government, (1794-6) "Minutes of the evidence taken before the Committee of Council appointed for the consideration of all matters relating to trade and foreign plantations… the 27th of April 1795 to the 5th of November 1795," BPP House of Lords Sessional Papers (19th November 1794 to 14th May 1796), 1-206.

British Government, (1813-14) "Report from the Select Committee on petitions related to the Corn Laws of this kingdom: together with the minutes of evidence and an appendix of accounts," BPP 1813-14, vol. 3, pp. 195-342.

British Government, (1821) "Report from the Select Committee to whom the several petitions complaining of the depressed state of agriculture of the United Kingdom, were referred," BPP 1821, vol. 9, pp. 1-479.

British Government, (1826-7) "Report from the Select Committee of the House of Lords appointed to inquire into the price at which foreign grain may be shipped in foreign ports; the quantity of such grain; and the price at which such grain can be imported into this country; and to report to the House together with the minutes of evidence taken before the said committee," BPP 1826-7, vol. 6, 633-848.

British Government, (1834) "Report from Select Committee on the sale of corn," BPP 1834, vol. 7, 1-532.

British Government, (1834) "Returns from corn inspectors of the customs or practices of selling corn in their districts," BPP 1834, vol. 105, 251-317.

British Government, (1852-3) *Census of Great Britain, 1851: Population Tables II*, BPP 1852-3, vol. 88, 290-3.

British Government, (1887) "Wheat and flour imports. Return of the number of quarters of wheat and wheat flour imported into the United Kingdom during the twenty years 1866 to 1886, giving the number of quarters imported from foreign countries and Brtish possesions in separate totals," BPP 1887, vol. 75, 299-304.

Burrell, Sir Charles (1841) "On some varieties of wheat," *Journal of the Royal Agricultural Society of England*, vol. 3, 147.

Capper, Charles (1862) *The port and trade of London* (London: Smith, Elder & Co).

Dodd, Frank (1856) *The food of London* (London: Longman, Brown, Green and Longmans).

Elliott, Sir T.H. (1903) Cabinet Memorandum.

Ellis, William (1744) *Agriculture improved* (London: T. Osborne).

Ellis, William (1745) *Chiltern and Vale farming explained* (London: T. Osborne).

Hall, Sir A. D. (1919) *The book of the Rothamsted experiments* (London: John Murray).

Handley, H. (1843) "Report on prize wheats," *Journal of the Royal Agricultural Society of England*, vol. 5, 397-8.

Hillyard, C. (1840) "On the productiveness of certain wheats," *Journal of the Royal Agricultural Society of England*, vol. 2, 65-6.

Humphries, A. E., and R. H. Biffen (1907-8) "The improvement of English wheat", *Journal of Agricultural Science*, vol. 2(1), 1-16.

Jago, William (1886) *Chemistry of wheat, flour and bread and technology of breadmaking* (Brighton and London: Simpkin, Marshall, Hamilton, Kent and Co).

Jago, William (1895) *Science and art of breadmaking: chemistry and analysis of wheat, flour, etc.* (Brighton and London: Simpkin, Marshall, Hamilton, Kent and Co).

Jacob, William (1834) Evidence in British Government, "Select committee on the sale of corn," BPP 1834, vol. 7, 411.

Johnson, George (1902) *Grain, flour, hay, and seed merchants' accounts* (London: Gee & Co.).

Lawes, J. Bennet, and J. H. Gilbert (1864) "Report of experiments on the growth of wheat for 20 years in succession on the same land," *Journal of the Royal Agricultural Society of England*, vol. 25, 93-185.

Lawes, J. Bennet, and J. H. Gilbert (1884) "On the continuous growth of wheat on the experimental plots at Rothamsted during the 20 years, 1864-83, inclusive," *Journal of the Royal Agricultural Society of England*, 2nd series, vol. 20, 391-481.

Lawes, J. Bennet, (1895) *Memoranda of the origin, plan and results of the field and other experiments conducted on the farm and in the laboratory of Sir John Bennet Lawes at Rothamsted, Hertfordshire* (Rothamsted: Sir John Benet Lawes).

Le Couteur, Colonel (1839) "On pure and improved varieties of wheat lately introduced into England," *Journal of the Royal Agricultural Society of England*, vol. 1, 113-23.

Loft, W. (1848) "On different varieties of wheat, and the advantage of thick sowing," *Journal of the Royal Agricultural Society of England*, vol. 10, 281-3.

Masters, M. T. (1888) "Wheat," in *Encyclopedia Britannica* (9th ed.; New York: Charles Scribner's & sons).

Miles, W. (1843) "Report on the wheats selected for trial at Cambridge, and on other wheats," *Journal of the Royal Agricultural Society of England*, vol. 5, 391-5.

Miles, W. (1846) "Report on the wheat selected for trial at Southampton and on other wheat," *Journal of the Royal Agricultural Society of England*, vol. 8, 566-72.

Page, Richard (1834) Evidence to British Government, "Select committee on the sale of corn," BPP 1834, vol. 7, 554.

Palmer, Nathaniel (1834) Evidence to British Government, "Select committee on the sale of corn," BPP 1834, vol. 7, 256-60.

Pasley, Colonel Charles W. (1834) Evidence in British Government, "Select committee on the sale of corn," BPP (1834), vol. 7, 273-9.

Percival, John (1934) *Wheat in Great Britain* (London: Duckworth).

Schafer, E. G., E. F. Gaines and O. E. Barbee (1921) *Wheat production: as influenced by variety, time of seeding, and source of seed* (Pullman, WA.: State College of Washington Agricultural Experiment Station).

Shelley, J. V. (1844) "Reports on wheats selected for trial at the Bristol meeting," *Journal of the Royal Agricultural Society of England*, vol. 6, 584-5.

Stead, Patrick (1834) Evidence in British Government, "Select committee on the sale of corn," BPP 1834, vol. 7, 95-6.

Trowell, Samuel (1750) *The farmer's instructor* (London: J. Hodges).

Weeks, James (1871) *The corn exchange ready-reckoner* (London: Alfred Boot).

Winter, George (1798) *A compendious system of husbandry* (London: E. Newberry).

Secondary sources.

Barquín, Rafael, (2005) "The demand elasticity for wheat in the 14th to 18th centuries," *Revista de Historia Económia*, vol. 23(2), 241-67.

Baten, Jörg, Dorothee Crayen and Hans-Joachim Voth, (2014) "Numeracy and the impact of high food prices in industrializing Britain, 1780-1850," *Review of Economics and Statistics*, vol. 6(3), 418-30.

Bennett, Richard and John Elton, (1898) *History of corn milling* (London: Simpkin, Marshall and Co).

Bourke, P. M. Austin (1976) "The Irish grain trade, 1839-48," *Irish Historical Studies*, vol. 20(78), 156-69.

Brunt, Liam, and Edmund Cannon (1999) "A Grain of Truth in Medieval Interest Rates? Re-examining the McCloskey-Nash Hypothesis" University of Bristol Economics Department Discussion Paper 98/462.

Brunt, Liam, and Edmund Cannon (2004) "The Irish grain trade from the Famine to the First World War," *Economic History Review*, vol. 57(1), 33-79.

Brunt, Liam and Edmund Cannon (2013) "The truth, the whole truth and nothing but the truth: the English Corn Returns as a data source in economic history, 1770-1914," *European Review of Economic History*, vol. 17(3), 318-39.

Collins, E. J. T. (1993) "Why wheat? Choice of food grains in Europe in the nineteenth and twentieth centuries," *Journal of European Economic History*, vol. 22(1), 7-38.

Clark, Gregory (1999) “Markets and economic growth: the grain market of mediaeval England,” unpub. mimeo., University of California-Davis.

Clark, Gregory (2001) “Farm wages and living standards in the Industrial Revolution: England, 1670-1869,” *Economic History Review,* vol. 54(3), 477-505.

El-Husseini, Farid (2002) *Interspatial price equilibrium and the incidence of tariffs*, University of Bristol Ph.D. dissertation.

Federico, Giovanni (2008) “The first European grain invasion: a study in the integration of the European market,” EUI Working Paper HEC 2008/1.

Feinstein, Charles H. (1998) “Pessimism perpetuated: real wages and the standard of living in Britain during and after the Industrial Revolution,” *Journal of Economic History*, vol. 58(3), 625-68

Fogel, Robert W. (1992) “Second thoughts on the European escape from hunger: famines, chronic malnutrition, and mortality rates,” in S. R. Osmani (ed.), *Nutrition and Poverty* (Oxford: Clarendon Press), 243-86.

Hassler, Uwe, and Vladimir Kuzin (2009) “Cointegration analysis under measurement errors,” in J. Binner, D. Edgerton and T. Elger (eds.) *Advances in Econometrics*, vol. 24, 131–150 (Bingley: Emerald Group Publishing Limited).

Henry, R. J., and P. S. Kettlewell (1996) *Cereal grain quality* (London: Chapman & Hall).

Hill, Lowell D. (1990) *Grain grades and standards* (Chicago: University of Illinois Press).

Hynes, William, David S. Jacks and Kevin H. O’Rourke (2012) “Commodity market disintegration in the interwar period,” *European Review of Economic History*, vol. 16(2), 119-43.

Jacks, David S. (2011) “Foreign wars, domestic markets: England, 1793-1815,” *European Review of Economic History*, vol. 15(2), 277-311.

Keller, Wolfgang, and Carol H. Shiue (2014) “Endogenous formation of free trade agreements: evidence from the Zollverein's impact on market integration,” *Journal of Economic History*, 74(4), 1168 - 1204.

McCloskey, Donald N., and John Nash (1984) “Corn at interest: the extent and cost of grain storage in medieval England,” *American Economic Review*, 74(1), 174-87.

Nielsen, Heino Bohn (forthcoming) “The Co-Integrated Vector Autoregression with Errors-in-Variables,” *Econometric Reviews*.

Olmstead, Alan L. and Paul W. Rhode (2002) “The Red Queen and the Hard Reds: productivity growth in American wheat, 1800–1940,” *Journal of Economic History*, vol. 62(4), 929-66.

Olmstead, Alan L. and Paul W. Rhode (2003) “Hog-round marketing, seed quality, and Government policy: institutional change in U.S. cotton production, 1920–1960,” *Journal of Economic History*, vol. 63(2), 447-88.

Persson, Karl Gunnar (2000) *Grain markets in Europe, 1500-1900* (Cambridge: University Press).

Persson, Karl Gunnar (2004) “Mind the gap! Transport costs and price convergence in the nineteenth century Atlantic economy,” *European Review of Economic History*, vol. (2), 125-47.

Petersen, Christian (1995) *Bread and the British economy* (Aldershot: Scolar Press).

Phelps Brown, E. H., and Sheila V. Hopkins (1956) “Seven centuries of the prices of consumables, compared with builders’ wage-rates,” *Economica*, 23(2), 296-314 .

Phillips, A. D. M. (1989) *Underdraining of farmland in England during the nineteenth century* (Cambridge: University Press).

Shiue, Carol H., and Wolfgang Keller (2007) “Markets in China and Europe on the eve of the Industrial Revolution,” *American Economic Review*, vol. 97(4), 1189-1216.

Solar, Peter M. (1987) *Growth and distribution in Irish agriculture before the Famine*, Stanford University Ph.D. dissertation.

Velkar, Aashish (2012) *Markets and measurements in nineteenth-century Britain* (Cambridge: University Press).

Walton, John R. (1999) “Varietal innovation and the competitiveness of the British cereals sector, 1760-1930,” *Agricultural History Review*, vol. 47(1), 29-57.

Figures and Tables

Figure 1. Bristol weekly wheat prices, 1790.

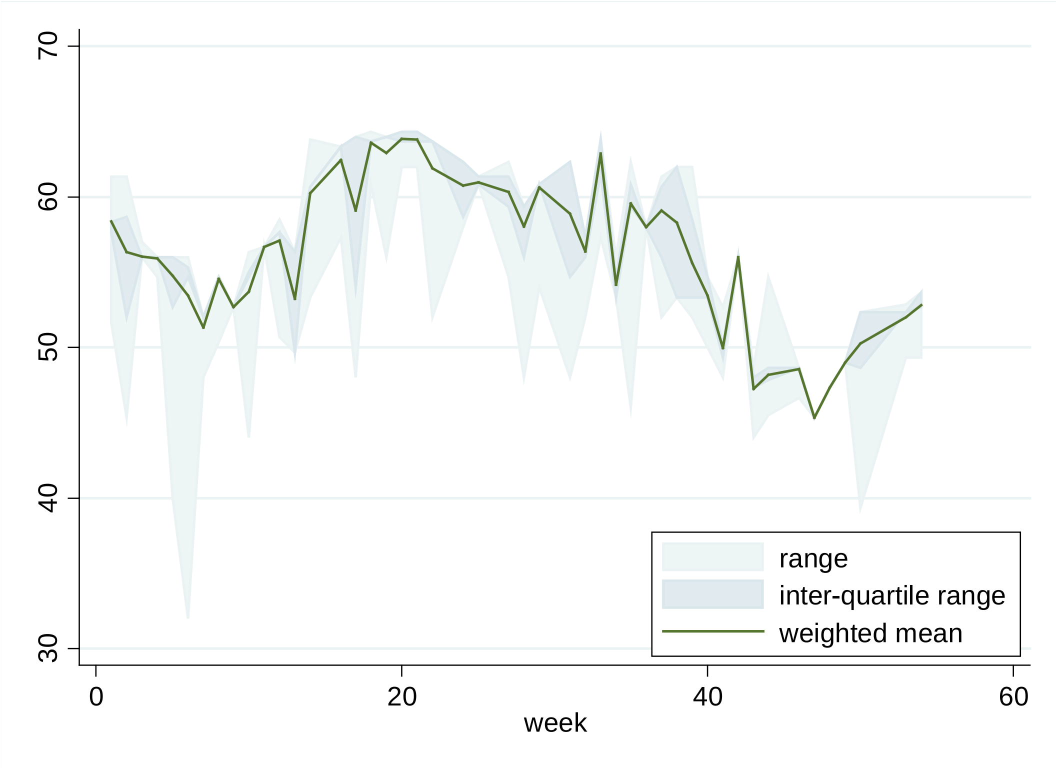


Table 1: Summary of information from Corn Inspectors in 1834

	Average, usual or mid-range bushel weight	Range (maximum - minimum) of bushel weights	Range of bushel weights relative to average
No of observations	46	18	18
Average	62.10	3.95	0.06
Minimum	58.63	0.50	0.01
Maximum	67.50	8.00	0.13

Source: authors' calculations from BPP (1834), 105; details in appendix A2.

Table 2: Relationship between price and bushel weight.

Dependent variable: ln(average price per bushel)	(1)	(2)	(3)
ln(bushel weight)	0.6108^{*}	0.6528^{+}	0.5527
	(0.2564)	(0.3515)	(0.3504)
Region FEs		✓	
County FEs			✓
N	40	40	40
Adjusted r²	0.1287	0.1129	0.4775

Notes. Regressions include a constant. SEs in parentheses, clustered at county level in specifications (1) and (2). Dependent variable is average log price, 1828-42, where markets are dropped if more than 5% of prices are missing (details in appendix). $^{+}$ $p < 0.10$, * $p < 0.05$.

Table 3. Relationship of price to bushel weight and other qualities.

	ln(price/bushel)	ln(price/bushel)
ln(bushel weight)	2.0806^{**}	1.0974^{+}
	(0.4205)	(0.4650)
ln(impurities)		-0.0079^{+}
		(0.0037)
ln(gluten)		0.0042^{+}
		(0.0019)
ln(flour)		0.0001
		(0.0040)
N	10	10
Adjusted r^2	0.7229	0.8763

Regressions also include a constant. Standard errors in parentheses $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 4: Bushel weights of British grain (lbs/Imperial bushel).

	Corn Returns Act 1791	Dodd 1856	Corn Returns Bill 1881	Brunt and Cannon 1839-1915	Johnson 1902
Wheat	59	60	60	60.75	63
Barley	51	48	50	50	56
Oats	39	40	38	38.8	42

Notes and sources. See appendix A4.C.

Table 5. Variations in the weight of an English bushel of wheat.

Evidence of Richard Page			
	High quality		Low quality
Good year	66		52
Bad year	60		48
Evidence of Patrick Stead			
	High quality	Second quality	Third quality
Good year	64	62	59
Bad year	60	57	55
Evidence of Colonel Charles Pasley			
Average	65		55

Notes. Evidence to Select Committee on the Sale of Corn, BPP (1834), 354, 95-6 and 277. Colonel Pasley, Royal Engineers, conducted a series of experiments in grain crops measurement methods. He also gives bushel weight variation for other crops: rye (51-58lbs); barley (46-51); oats (36-44); peas (61-69); beans (61-68).

Table 6. Flour produced by wheat of various bushel weights.

Bushel weight	Flour content by category (%, by weight)							
lbs/bu	House-holds	Second quality	Third quality	Fine middlings	Coarse middlings	Flour total	Pollard, bran	Waste
Panel A. *c*. 1788								
58	54.5	16.3	6.8	4.0	0	81.6	17.1	1.3
Panel B. *c*. 1795								
62	64.5	6.7	0.0	4.4	3.0	78.6	19.4	2.0
60	65.4	6.5	0.0	4.4	2.9	79.2	18.8	2.1
58	61.6	6.0	0.0	4.7	3.9	76.2	21.1	2.6
56	0.0	61.8	4.5	4.2	4.0	74.5	22.8	2.7
54	0.0	54.5	6.0	6.5	6.5	73.5	23.7	2.8
Panel C. *c*. 1841								
64						82.1		
63						81.3		
62						79.9		
61						76.7		
Panel D. *c*. 1856								
63	77.78			2.0	1.6	81.4	16.5	2.2

Notes. Calculations in appendix A4.D. Panel A based on Albion Mill data (Bennett and Elton, *History of corn milling*, vol. 3, 290); Panel B from evidence of Robert Ardlie, Appendix 6 of British Government, "Fourth report"; Panel C on Miles, "Report… Cambridge", 391-5 and Le Couteur, "On the pure and improved varieties", 113-23; Panel D from Dodd, *Food*, 184-5.

Table 7. Relationship between wheat bushel weights (lbs) and wheat prices (d)

	First quality grain	Second quality grain	Third quality grain
Good harvest			
Bushel weight	64	62	59
Price (d/bu)	480	456	420
Price (d/lb)	7.50	7.35	7.12
Bad harvest			
Bushel weight	60	57	55
Price (d/bu)	960	888	840
Price (d/lb)	16.00	15.58	15.27

Source. Evidence of Patrick Stead, British Government, “Select committee,” BPP 1834, 95-6.

Table 8. Yields of wheat cultivars in bu/acre (bushel weights in brackets).

Cultivar	*c.* 1841	*c.* 1871	*c.* 1914
Club Wheat [Square Head]		49.2 (61.1)	38.7 (61.7)
Red Rostock [Russian Red]		45.4 (59.9)	36.6 (59.5)
Red Chaff		39.0 (61.5)	37.4 (60.4)
White Chiddam	22.5 (62.7)	37.1 (62.0)	
Golden Drop	27.6 (61.3)	46.8 (62.5)	
Old Red Lammas	30.5 (63.2)	39.6 (62.6)	
White Chaff	44.9 (59.8)	48.9 (61.0)	
Bole's Prolific	49.8 (61.3)	44.0 (61.5)	
Bristol Red	45.6 (62.0)	42.1 (61.3)	
MEAN	36.8 (61.7)	43.1 (61.8)	

Notes and sources. We matched all possible cultivars, based on sources cited in table 4; 1841 estimates based on very small N – Old Red Lammas has four observations, other cultivars only two!

Table 9. Imported wheat bushel weights (lbs), *c.* 1825 and *c.* 1900.

Panel A: c. 1825						
Britain	Ireland	Saarbrücke	Holstein	Danzig	Odessa	Taganrog
60.75	60	56	57.5	63	59	70
Panel B: c. 1900						
Britain	Australia	Argentina	India	Russia	USA	
63	63	62	62	62.5	63	

Notes. Panel A: British Government, "Report", BPP 1826-7, 683 (Odessa), 700 (Holstein, Danzig), 725 (Taganrog); Brunt and Cannon (2004, 35-6) (Britain); British Government, "Report", BPP 1821, 307 (Ireland), 371 (Saarbrücke). Panel B: Johnson (1902, 5,20,22).

Table 10. Averages prices of wheat traded in England, 1855-74 (d/bu).

	Mean price, 1855-63	Import share, 1855-63	Mean price, 1864-74	Import share, 1864-74
English	85.14		81.77	
American	85.33	23.01	81.44	23.76
Canadian	83.12	4.44	79.66	4.51
Chilean			84.35	2.26
Australian			96.33**	1.01
Prussian	88.98**	14.06	86.66**	9.54
Mecklenburgian	87.65	1.98		
Danish	79.98**	3.93	77.64*	1.05
Belgian	84.22	0.55		
Dutch	83.81	0.56		
French	80.52	5.46	76.12*	2.62
Austrian Italy	85.28	1.13		
Russian	73.37**	14.08	74.48**	24.66
Romanian	70.55**	1.28	70.91**	0.97
Turkish	72.19**	1.51	71.16**	1.86
Egyptian	57.79**	7.06	70.93**	2.36
Total		79.02		74.60

Notes and sources. British Government, "Annual trade and navigation accounts of the United Kingdom", various years. Data begin only in 1855 and format changes from 1864 onwards, as discussed below. We include all countries exporting to England every year (or nearly so) – important because volatility is high, so including occasional exporters could generate erratic results. German prices for Prussia from 1870 onwards. Import share is share of *total quantity imported from all sources*, 1855-63 and 1864-74. * and ** means prices significantly different from English at 5% and 1% levels (matched pairs two-tailed t-test).

Figure 2. European/English wheat price relatives (all traded in London).

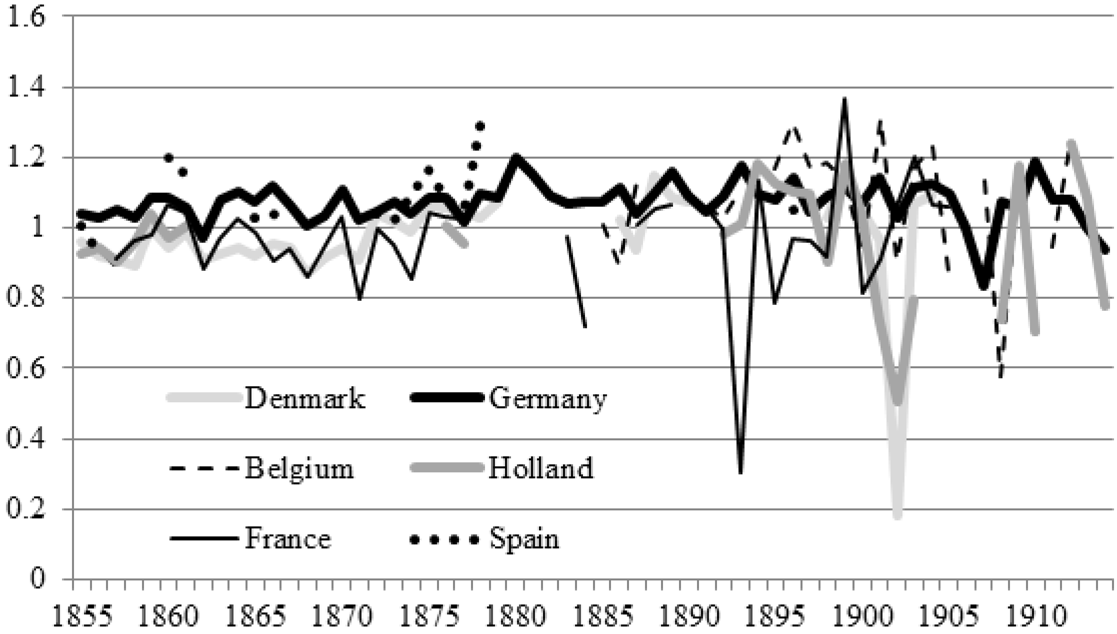


Figure 3. New World/English wheat price relatives (all traded in London).

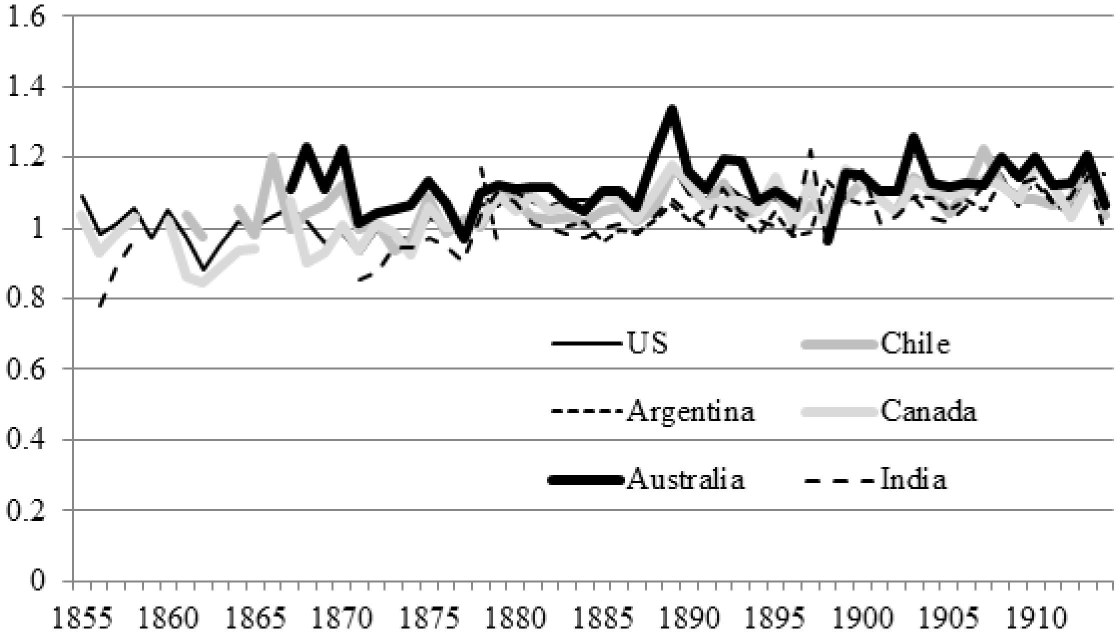


Figure 4. Eastern Mediterranean/English wheat price relatives (traded in London).

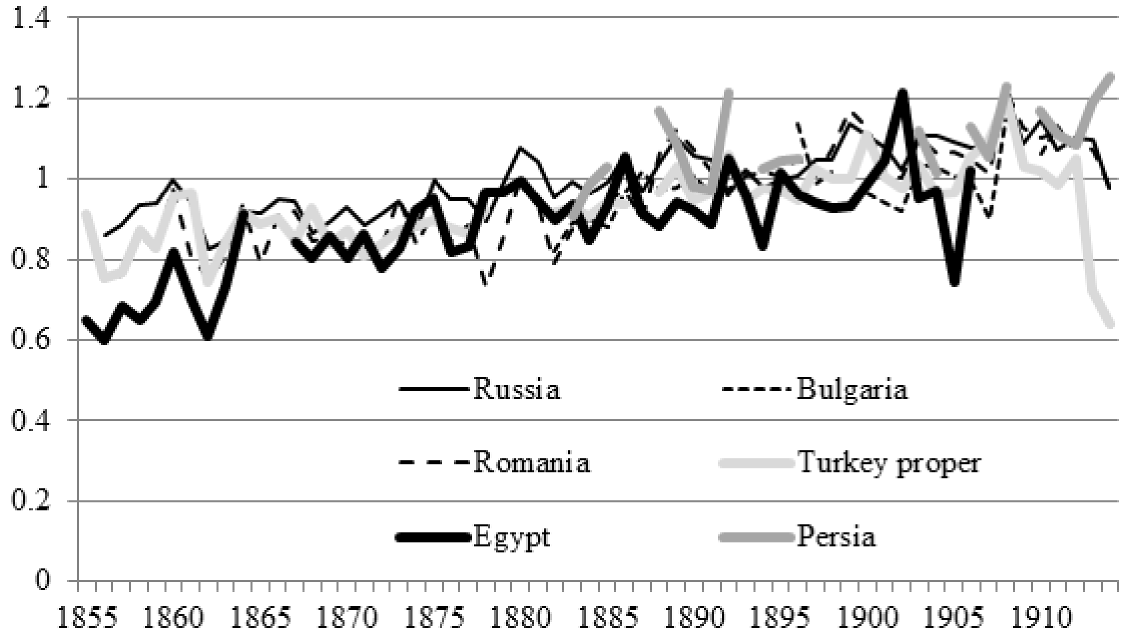


Figure 5. Prices of English wheat, wheat consumed in England and differentials.

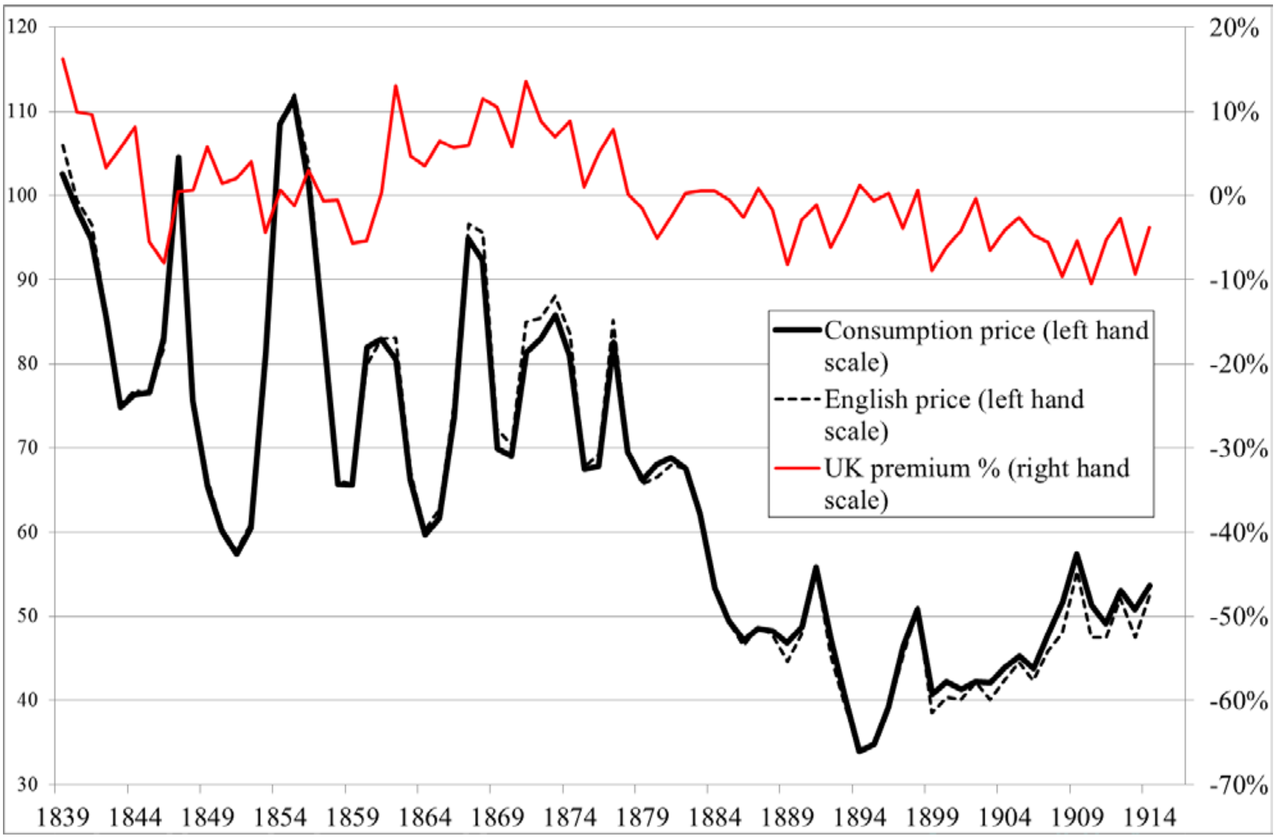


Table 11. Ratio of quality variation to price variation.

$\sigma_\eta / \sigma_\varepsilon$	0	0.25	0.50	1.00
σ_η / σ_p	0	0.23	0.40	0.58

Figure 6. Asymmetric loadings: estimated effect of shock to price A

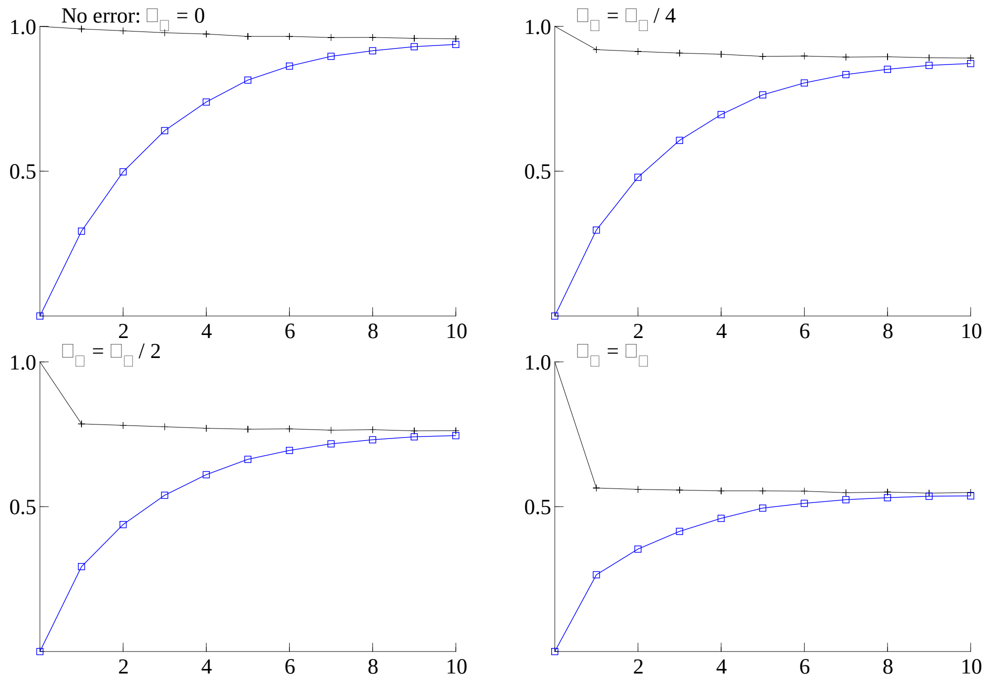


Figure 7. Asymmetric loadings: estimated effect of shock to price B

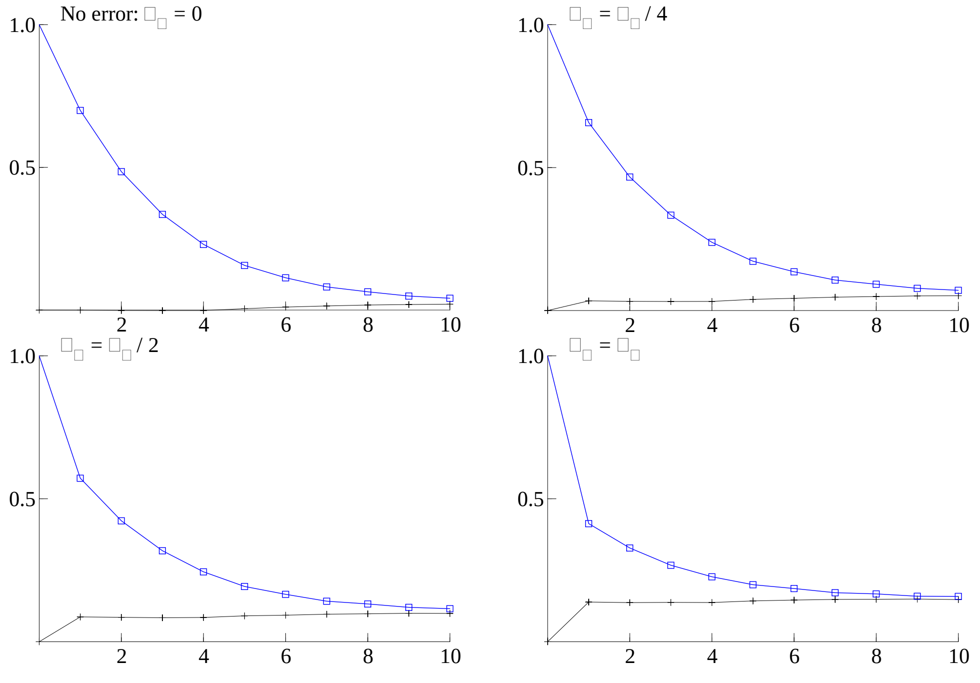


Figure 8. Symmetric loadings: estimated effect of shock to one price

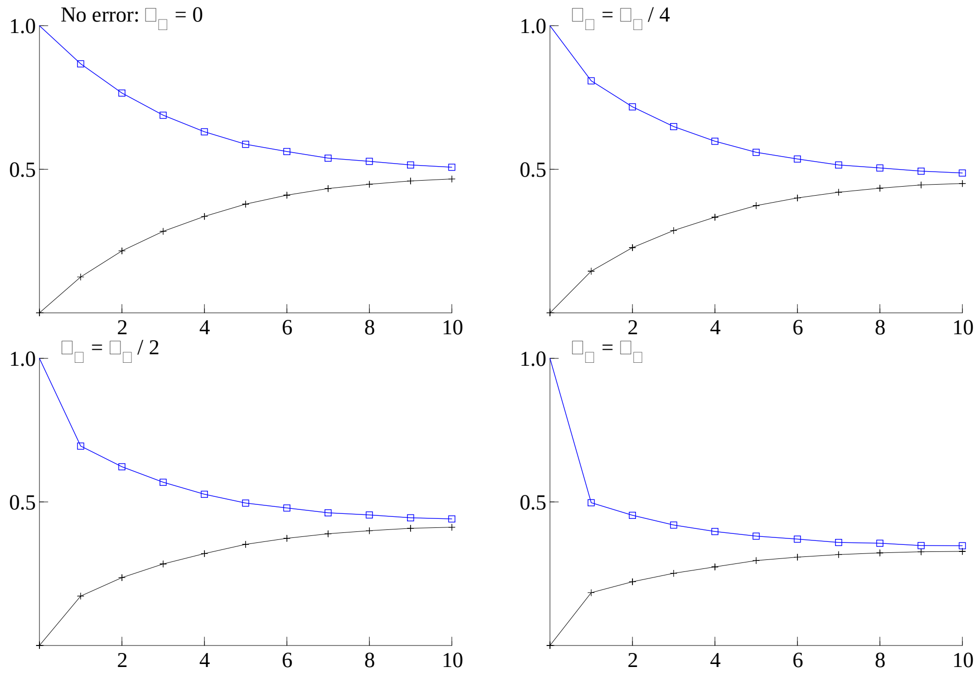


Table 12. Effect of measurement error on half life estimates.

True Half Life	Estimated half life as proportion of true half life			
	$\sigma_\eta = 0$	$\sigma_\eta / \sigma_\varepsilon = 1/4$	$\sigma_\eta / \sigma_\varepsilon = 1/2$	$\sigma_\eta / \sigma_\varepsilon = 1$
1	98%	89%	75%	61%
2	97%	83%	49%	36%
3	95%	80%	49%	25%
4	93%	78%	48%	20%
5	90%	77%	47%	16%
10	79%	67%	43%	9%

Table 13. Relationship between ln(bushelweight) and ln(yield).

Panel A. Time series data for two plots				
	(1) manured 1844-83	(2) manured sub-sample	(3) pooled 1844-83	(4) pooled sub-sample
ln(yield)	0.0838*	0.0365	0.0559***	0.0366***
	(0.0362)	(0.0354)	(0.0139)	(0.0062)
N	40	25	80	50
r^2	0.177	0.0330	0.261	0.182
Panel B. Panel data for different cultivars				
	(5) between-group estimator 1871-81	(6) between-group estimator 1871-78	(7) within-group estimator 1871-81	(8) within-group estimator 1871-78
ln(yield)	0.0985**	-0.0342	0.114***	0.0532***
	(0.0315)	(0.0279)	(0.0088)	(0.0084)
N	239	171	239	171
r2	0.289	0.0613	0.604	0.121

Notes. Regressions include unreported constants; robust standard errors in parentheses: regression (3) and (4) standard errors clustered by year. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendices.

A1. Creating a US harvest quality index

The Chicago Board of Trade laid down explicit guidelines for grading wheat, and it is worth considering the grading rubric. Grade 1 winter wheat had to be "sound, plump and well cleaned"; Grade 2 had to be "sound and reasonably clean"; Grade 3 was "not clean or plump enough for Grade 2, but weighing not less than 54lbs per measured bushel"; Grade 4 included wheat that was "damp, musty, or from any cause so badly damaged as to render it unfit for No. 3"; and "No grade" was inferior to that (*Chicago Board of Trade Annual Report 1888*, 94; essentially the same regulations were printed in every annual report). The guideline wording varied slightly from one wheat type to another, but was essentially the same for all types. Wording also changed slightly over time; for example, "No Grade" was initially known as "Rejected" and later became "Standard Grade"; we refer to it throughout as "No Grade".

Note further the legal wheat bushel weight in most of the US was 60lbs (*Chicago Board of Trade Annual Report 1888*, 124), implying Grade 1 and 2 wheats weighed at least that much. It is specifically stipulated Grade 3 wheats (and, most likely, Grade 4 wheats) must weigh 54lbs minimum. So Grade 3 and 4 wheats should trade at a 10% discount to Grades 1 and 2, simply because Grade 3 was less dense by around 10% (=6/60) and generated at least 10% less flour.

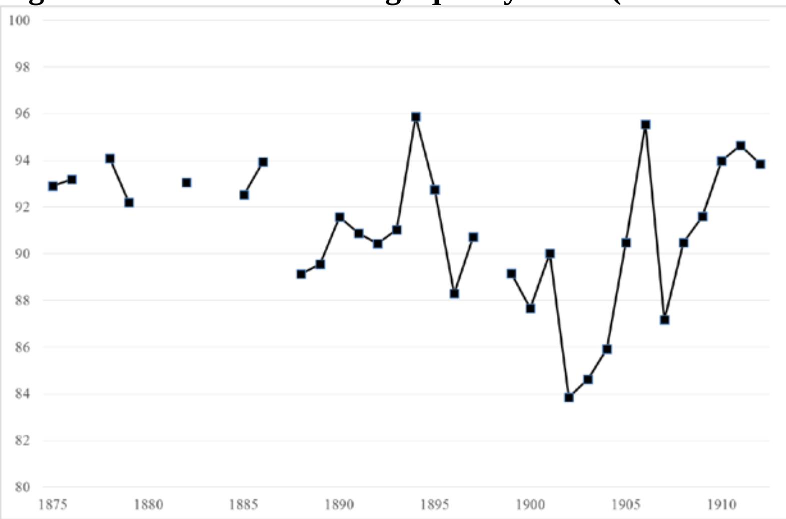
Consider price differentials for different qualities of the same wheat type. The Chicago Board produced weekly price quotations for various qualities of wheat in store. We do not have data on all types in all years – only a snapshot for each year, where the snapshots capture slightly different things each time – so our data are illustrative rather than definitive. However, they are quite informative. Take the 1876 data for Spring Wheat (*Chicago Board of Trade Annual Report 1876*, 81). Grade 1 traded at a 5% premium over Grade 2; Grade 3 at a 12% discount to Grade 2; and No Grade at a 25% discount to Grade 2. Since a Grade 3 bushel weighed around 10% less than a Grade 2 bushel, a 12% discount seems reasonable. In 1885 (*Chicago Board of Trade Annual Report 1885*, 113), Grade 4 Spring Wheat traded at a 29% discount to Grade 2.

We need to estimate US harvest quality time series volatility. Chicago Board of Trade data permit this, providing data on total wheat rail cars numbers inspected annually, disaggregated by type and grade. First, create shares of each wheat (by type and grade). Since the price cross section in a given year reflects quality variation, we can weight each grain type and grade by its relative price. In years with relatively plentiful high quality types and grades, our price-weighted quality index will be high, and *vice versa*. But relative prices in a given year are influenced by relative availability of different types and grades: for example, prices of high qualities are relatively lower in years with relatively much high quality grain (a kind of simultaneity bias). We can avoid this problem using a fixed basket of price weights (rather like a Laspeyres index). This is the approach that we follow.

As noted above, we have nothing like an exhaustive set of prices for all wheats (all five qualities of every type). So we first aggregate across all types, by grade, to get an overall quality distribution for each year (Grade 1 down to No Grade). We then take illustrative relative prices of Spring Wheat in 1876/85 and apply them to annual wheat grade distributions. This generates an annual harvest quality index, from which we calculate time series volatility. Note this is an underestimate of volatility because we abstract from type fluctuations within each grade: some years have more high-value Red Winter in Grade 3, and less low-value Hard Winter, and so on. By assuming identical price relatives for all Grade 3 types, we assume away this source of volatility.

The quality index coefficient of variation is 3.3%, 1875-1912 (the years for which we found data), similar to England around the same time. Figure A1.1 charts the index: if all wheat were second-quality then the index would equal 100.

Figure A1.1 US wheat average quality index (Grade 2 = 100)



A2. Implications of unobserved quality

Here we characterize the consequences of unobserved quality (interpreted as a form of measurement error). The main text decomposes price per volumetric measure as:

$$\underbrace{P^V}_{\substack{\text{shillings}\\\text{per bushel}}} = \underbrace{P^*}_{\substack{\text{shillings per}\\\text{quality-adjusted}\\\text{lb}}} \times \underbrace{\theta}_{\substack{\text{quality}\\\text{adjustment}}} \times \underbrace{B}_{\substack{\text{lb per}\\\text{bushel}}} \Leftrightarrow \ln P^V = \ln P^* + \ln\theta + \ln B \qquad \text{(A.1)}$$

where P^V is price by volume (typically measured in shillings per bushel in our data); B is grain density (typically the bushel weight in our data); θ is any remaining quality adjustment orthogonal to density; and P^* is price of constant-quality grain by mass. Since θ is a residual, A.1 is true by definition. Variations in P^* are due to supply and demand changes (attenuated by inter-period storage and inter-market trade).

Consequences of A.1 depend on the precise question asked; here we consider several possibilities.

A2.1 Different prices in the same market at the same point in time.

Any variation in a data set comprising prices in the same market at the same time must be due to market failure or quality variation. Suppose it is the latter. Note that P^* is a constant (there is only one market), so $\text{var}\left[\ln P^*\right] = \text{cov}\left[\ln P^*, \ln\theta + \ln B\right] = 0$. Equation A.1 can be written more precisely as:

$$\ln P_i^V = \ln P^* + \ln\theta_i + \ln B_i \qquad \text{(A.2)}$$

where different prices are sub-scripted with i. Price variation is due to variation in B_i and θ_i. Regressing $\ln P_i^V$ on $\ln B_i$ yields an unbiased unit parameter estimate only if $\text{cov}\left[\ln\theta, \ln B\right] = 0$ (when the unobserved regression disturbance is simply $\ln\theta_i$). The r-squared from such a regression depends on bushel weight variance and the relationship between bushel weight and the remaining quality denoted in θ_i.

A2.2 Different prices in different markets at the same point in time

Take a price cross-section from different markets (measured at the same time), so that A.1 can be written more precisely as:

$$\ln P_i^V = \ln P_i^* + \ln\theta_i + \ln B_i \qquad \text{(A.3)}$$

Unless the Law of One Price (LOOP) is satisfied perfectly, $\mathrm{var}\left[\ln P^{*}\right] \neq 0$. Regressing observed volumetric price on bushel weight, the disturbance is $\ln P_{i}^{*} + \ln \theta_{i}$; whether or not the expected estimate of the bushel weight coefficient is unity depends on covariance of bushelweight with this disturbance.

However, there is a further complication. Section A2.1 considered the possibility of having actual prices of individual trades (similar to data illustrated in figure 1). But researchers typically compare prices in different markets using data on average prices in those markets; so A.1 is specified more correctly as:

(A.4) $$\ln \overline{P}_{j}^{V} = \ln P_{j}^{*} + \overline{\ln \theta_{j}} + \overline{\ln B_{j}}$$

where a bar over a variable indicates an average within a market and sub-script j refers to the market (rather than an individual price). A consequence is that the r-squared from regression estimates of A.4 may be higher or lower than A.2, because we are unable to say which of the two disturbances has a higher variance.

A2.3 Difference prices in the same market at different points in time

Data constraints preclude us from regressing price on bushel weight over time. Such regressions would be subject to similar issues to those discussed in the previous sub-section. Time series data are more commonly used to estimate demand curves or price relationships between different markets. Suppose we had annual data on average prices so the precise way of writing A.1 were:

(A.5) $$\ln P_{t}^{V} = \ln P_{t}^{*} + \ln \theta_{t} + \ln B_{t}$$

where t denotes year. Inter-year variation is one of the biggest in historical prices due to harvest volatility: good harvests result in higher yields and lower underlying prices $\ln P_{t}^{*}$. The yield-price relationship depends partly on elasticity of demand; combined with the section 6 result that bushel weight depends on yield, it follows $\mathrm{cov}\left[\ln P_{t}^{*}, \ln B_{t}\right] \neq 0$. Not only is there measurement error, but it need not be "classical measurement error". For example, if observed price $\ln P_{t}^{V}$ were used as an explanatory variable, then the resulting regression disturbance would be correlated with the unobserved price $\ln P_t^*$, creating a different bias from when the disturbance is correlated only with observed price.

A3. Summary of returns from Corn Inspectors

The 148 England and Wales Corn Inspectors in 1834 were asked to give information on markets they monitored. (Notionally, 150 markets were monitored but two markets – Windsor and Beccles – had no inspector at that time). Corn Inspectors' replies were published in BPP (1834), vol. 105, 251-317. We entered data described in table A2.

Table A2. Data entered from the Corn Inspectors' 1834 report.

Town/market name (with the 1834 return spelling, and an alternative for ease of comparison)
The unit reported as being used for trade
Whether or not Winchester measure was still in use
Summary or quote of the text where we think it interesting
Name of the inspector
Average or customary price; where inspectors give a price range, the mid-point of the range
Upper and lower prices, the subsequent range, the range relative to the average price

Data are available in an Excel spreadsheet and Stata data file (with accompanying "do" file).

A4. Calculating relative standard deviations from regression analysis

Unadjusted r-squared is 0.153 in the first table 2 regression (adjusted r-squareds are reported in the table to facilitate comparison of regressions with different numbers of explanators).

Ignore theta in equation (3), so that:

(A.7) $$\ln P^V = \ln P^* + \ln B$$

Regression r-squared can be interpreted as:

(A.8) $$R^2 = \frac{\operatorname{var}\left[\ln B\right]}{\operatorname{var}\left[\ln P^V\right]} = \frac{\operatorname{var}\left[\ln B\right]}{\operatorname{var}\left[\ln P^*\right] + \operatorname{var}\left[\ln B\right]}$$

where $\ln P^*$ is the unobserved disturbance, and where $\operatorname{cov}\left[\ln P^*, \ln B\right] = 0$ (by construction) in an OLS regression. It follows:

$$\frac{\text{st.dev.}\left[\ln B_i\right]}{\text{st.dev.}\left[\ln P_i^*\right]} = \sqrt{\frac{R^2}{1-R^2}} = \sqrt{\frac{0.1534}{1-0.1534}} = 0.4257$$

These figures so far assume regression 1 in table 2 is the correct regression. The regression 3 partial r-squared is 0.079, suggesting a ratio of 0.293 instead of 0.426. Arguably, the third regression is superior because it controls more factors (county fixed effects). But there may be county-level quality variations orthogonal to bushel weight: for example, as discussed in section 3, Lincolnshire may grows Rivet more than other counties and Rivet is known to contain less gluten.

We have ignored quality variation not captured by bushel weight. Recall:

(A.9)
$$\begin{aligned}\operatorname{var}\left[\ln P^V\right] &= \operatorname{var}\left[\ln P^*\right] + \operatorname{var}\left[\ln\theta\right] + \operatorname{var}\left[\ln B\right] \\ &\quad +2\left\{\operatorname{cov}\left[\ln P^*, \ln\theta\right] + \operatorname{cov}\left[\ln P^*, \ln B\right] + \operatorname{cov}\left[\theta, \ln B\right]\right\}\end{aligned}$$

A5. Grain quality and bushel weight

The accompanying Stata file contains the regression analysis underlying table 2. Table A5.1 details elasticity estimates from Barclay (1845) and Hillyard (1840).

Table A5.1. Relationship between price and bushel weight from trials

Price per bushel (d)	Bushel weight	
	Barclay	Hillyard
65		84
64.75		84
64.5	97.5	
64.25		
64		81
63.75		81
63.5	94.5	
63.25		
63	95.25	
62.75		
62.5		78
62.25		
62	93	

Notes and sources. Barclay (1845, pp.192-3); Hillyard (1840, pp. 65-6). Elasticities calculated for the text are ln(97.5/93)/ln(64.5/62) and ln(84/78)/ln(65/62.5).

Table A5.2. Complete analysis of McDougall data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lpw	lpw	lpw	lpw	lpb	lpb	lpb	lpb
Bushel	0.010	0.0178^{+}	0.015	0.002	0.027°	0.035°	0.032°	0.018
weight	(0.008)	(0.007)	(0.009)	(0.008)	(0.009)	(0.007)	(0.009)	(0.008)
Indian			-0.071				-0.071	
			(0.033)				(0.032)	
Impurities			-0.006	-0.008			-0.006	-0.008
			(0.005)	(0.004)			(0.005)	(0.004)
Gluten			0.002	0.004			0.002	0.004
			(0.003)	(0.002)			(0.003)	(0.002)
Flour			-0.004	0.000			-0.004	0.000
			(0.005)	(0.004)			(0.005)	(0.004)
Constant	3.222*	2.801*	3.251*	3.738*	0.119	-0.310	0.157	0.639
	(0.522)	(0.427)	(0.584)	(0.426)	(0.523)	(0.430)	(0.580)	(0.425)
N	14	10	14	10	14	10	14	10
Adj. r^2	3.67	37.41	59.06	72.50	40.58	71.45	75.14	87.67

Standard errors in parentheses. $^{+}$ $p < 0.05$, $^{\circ}$ $p < 0.01$, * $p < 0.001$

Table A5.3. Nathaniel Palmer's wheat storage account.

Year	Wheat in store (quarters.)	Annual losses	Losses (%) [Palmer]	Losses (%) [Brunt-Cannon]
	(i)	(ii)	(iii)	(iv) = (ii) ÷ (i)
1814	19133	194.125	1.1	1.0
1815	53115	747.500	1.3	1.4
1816	73882	1766.000	2.3	2.4
1817	46036	593.125	1.3	1.3
1818	49100	816.750	1.5	1.7
1819	56357	1182.250	2.0	2.1
1820	11811	306.500	2.4	2.6
1821	10117	297.500	2.7	2.9
1822	12882	225.250	1.5	1.7
1823	15621	336.000	2.1	2.2
1824	13502	309.875	2.2	2.3
1825	33958	830.500	2.3	2.4
1826	31797	841.375	2.5	2.6
1827	21181	489.500	2.2	2.3
1828	32959	618.875	1.7	1.9
1829	78758	1426.875	1.7	1.8
1830	44259	866.375	1.7	2.0
1831	41086	1006.750	2.3	2.5
1832*	31430	750.125	2.3	2.4
1833*	14079	293.250	2.0	2.1
Total	691063	13898.500	2.0**	
		Simple average		2.1%
		Weighted average		2.0%
* If there be deducted from the years:		1832	4551.875	272.100
		1833	2060.750	392.500
		Total	6612.625	392.625

Hot Canadian shipments, the average loss % for the two years will be: 1832, 1¾; 1833, 1⅜; thus reducing the 20 years' average to 1 qr. 7 bus. 2 pks. 5 qts. per cent

Sources and notes. Columns (i)-(iii) from British Government, "Report," BPP (1834), 259. The original table measures losses in quarters and bushels, whereas we express them in quarters. Column (iv) - authors' calculations based on columns (i) and (ii), differing slightly from Palmer's calculations. ** the precise average given is 2 quarters 0 bushels 0 pecks. 2 quarters ≈ 2.016%

Table A5.4. Bushel weights of British grain (lbs/Imperial bushel).

	Corn Returns Act 1791	Dodd 1856	Corn Returns Bill 1881	Brunt and Cannon 1839-1915	Johnson 1902
Wheat	59	60	60	60.75	63
Barley	51	48	50	50	56
Oats	39	40	38	38.8	42
Rye	57	54		60	
Maize				59.75	60
Peas	64	60		63	
Beans	63	60			
Wheatmeal	49			50	
Barleymeal	48				
Ryemeal	53				
Beanmeal	48				
Oatmeal	22				

Notes. To 12th July 1827, the Winchester bushel was the most generally accepted measure; thereafter, use of the Imperial bushel – larger by the ratio of 32/31 (i.e. 3.2%) – was imposed by law. This table reports Imperial bushel weights to maintain consistency. But here we additional report official Winchester bushel weights, as laid down in the 1791 Corn Returns Act, to facilitate comparison across contemporary sources: wheat (57), barley (49), oats (38), rye (55). Naturally, we rebase any time series quantities or prices to Imperial bushels. Corn Returns Act data are from William Jacob, Comptroller of Corn Returns: British Government, "Select committee on the sale of corn," BPP (1834), vol. 7, 411, adopted also in Solar, *Growth*, 221; 1856 data from Dodd, *Food*, 184-5; 1881 data from the Corn Returns Bill BPP (1881, bill 17), which was primarily concerned with tithes and never passed; 1839-1915 data from Brunt and Cannon, "Irish grain trade", 35-6; 1902 data from Johnson, "Grain", 2.

Table A5.5. Flour produced by wheat of various bushel weights.

lbs/bu	lbs or %	House-holds	Second quality	Third quality	Fine middlings	Coarse middlings	Flour total	Pollards, bran	Waste
Panel A. *c*. 1788									
58	2320	1265.4	378	157.6	92.4	0	1893.4	395.6	30
58	100	54.5	16.3	6.8	4.0	0	81.6	17.1	1.3
Panel B. *c*. 1795									
62	496	320	33	0	22	15	390	96	10
60	480	314	31	0	21	14	380	90	10
58	464	286	28	0	22	18	354	98	12
56	448	0	277	20	19	18	334	102	12
54	432	0	235.5	26	28	28	317.5	102.5	12
62	100	64.5	6.7	0.0	4.4	3.0	78.6	19.4	2.0
60	100	65.4	6.5	0.0	4.4	2.9	79.2	18.8	2.1
58	100	61.6	6.0	0.0	4.7	3.9	76.2	21.1	2.6
56	100	0.0	61.8	4.5	4.2	4.0	74.5	22.8	2.7
54	100	0.0	54.5	6.0	6.5	6.5	73.5	23.7	2.8
Panel C. *c*. 1841									
64	512						420.3		
63	504						409.6		
62	496						396.1		
61	488						374.3		
64	100						82.1		
63	100						81.3		
62	100						79.9		
61	100						76.7		
Panel D. *c*. 1856									
63	504	392			10	8	410	83	11
63	100	77.78			2.0	1.6	81.4	16.5	2.2

Notes. Panel A based on Albion Mill data – Bennett and Elton, *History of corn milling*, vol. 3, 290; Panel B from Mr. Robert Ardlie, Appendix 6 of British Government, “Fourth report”; Panel C from Miles, “Report… Cambridge”, 391-5 and Le Couteur, “On the pure and improved varieties”, 113-23; Panel D from Dodd, *Food*, 184-5.

Table A5.6 Yields (bu/acre) and densities (lbs/bu) of different wheat cultivars.

Name of wheat	***c.* 1750**	***c.* 1841**		***c.* 1871-78**		**c. 1914-19**	
		Yield	Density	Yield	Density	Yield	Density
White Lammas	High yielding	37.0	61.6				
Red Lammas	Finest flour	30.5	63.2	39.6	62.7		
Yellow Lammas	High yielding; soft; liked by bakers						
Pickey	?						
Dame	?						
Smyrna/Turkey [Turkey Red]	?					37.3	62.9
White Cone	?						
Red Pirkey	?						
Yellow Pirkey	?						
White Pirkey	?						
Pirkey	?						
White-Brown	?						
Dugdale	Coarse, hardy; sown mostly in north; going out of fashion						
Spring Wheat	?						
White Wheat	?						
Rivet				53.6	58.7		
White Chaff [Square Head's master]		44.9	59.8	48.8	61.0		
Club Wheat [Square Head]				49.2	61.1	38.7	61.7
Golden Drop		27.6	61.3	46.8	62.3		
Bole's Prolific [Pilgrim's Prolific]		49.8	61.3	44.0	61.4		
Hardcastle				44.6	61.3		
Red Rostock [Russian Red]				45.4	59.9	36.6	59.5
Red Langham				41.6	61.7		
Bristol Red		45.6	62.0	42.1	61.3		
Red Wonder				42.3	61.3		
Red Chaff				39.0	61.5	37.4	60.4
Browick				41.7	61.0		
Casey's White				42.1	60.6		
Nursery Red				39.1	63.0		
Woolly Ear				41.3	61.3		
Golden Rough Chaff				40.4	62.0		
Chubb Wheat				41.2	60.8		
Original Red				36.6	59.3		
Victoria White				40.4	61.7		
White Chiddam		22.5	62.7	37.1	62.0		
Hunter's White				37.8	60.6		
Egyptian Mummy		45.0	62.0				
Essex Brown		40.0	64.0				

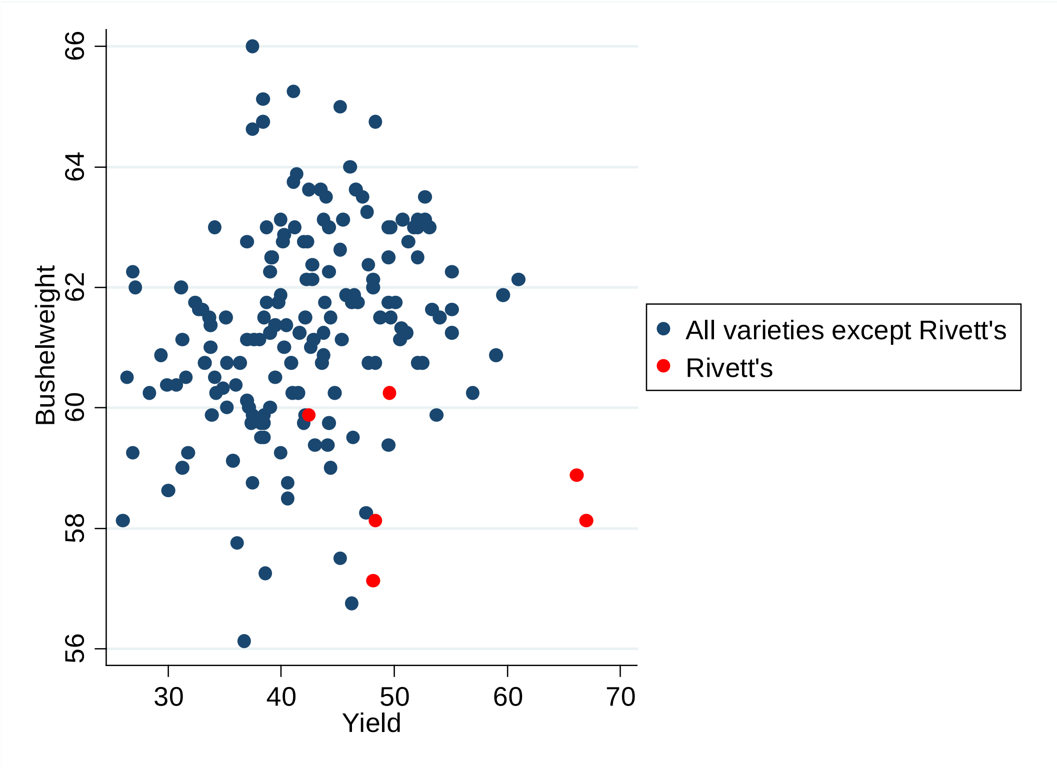
Name of wheat	***c.* 1750**	***c.* 1841**		***c.* 1871-78**		**c. 1914-19**	
		Yield	Density	Yield	Density	Yield	Density
Silver Drop		32.6	63.0				
Surrey White		36.0	64.0				
Red Straw White		35.7	58.5				
Brittany [Breedon]		30.9	59.8				
Spalding's		45.5	62.0				
London Red		39.8	63.3				
Red Cluster		47.8	63.0				
Syer's		46.5	61.0				
Soothy's		50.3	56.0				
Piper's Thickset		39.3	61.0				
Alfriston White		32.0	60.3				
Clover's		40.6	63.4				
Snowdrop White		39.0	63.0				
Whittington White		36.0	62.0				
Hopetown		41.4	60.4				
Golden Swan		32.8	63.7				
Red Champion		27.6	62.3				
Britannia		40.3	62.0				
Red Marigold		50.9	62.8				
Creeping Red		35.9	62.3				
Glory in the West		32.1	61.8				
Dantzig		30.2	61.3				
Salmon Brown		37.8	60.3				
Fenton		47.3	61.0				

Notes and sources: Varieties in *italic* are white wheats (have white kernels); varieties in ***bold italic*** are red wheats; the colour of other varieties is not known. Note some varieties have "Red" in the name but were definitely "white wheats" – such as *Red Chaff*. Names [in square brackets] are later names for the same wheat. A question mark ? indicates this cultivar was in recorded use, but nothing else is known. 1750 evidence from Ellis, *Agriculture*, 3-4, 33, 134; Ellis, *Chiltern and Vale farming*, 198, 207-9; Trowell, *Farmer's instructor*, 9; *c*. 1841 from Burrell, "On some varieties", 147; Handley, "Report", 397-8; Hillyard, "On the productiveness", 65-6; Loft, "On different varieties", 281-3; Miles, "Report… Cambridge", 391-5; Miles, "Report… Southampton", 566-72; and Shelley, "Reports", 584-5; *c*. 1871 from Lawes, *Memoranda*, 86-7, and is the 1871-8 average. The Rothamsted seed trial continued after 1878 but Lawes stated results were unreliable owing to technical cultivation problems; hence we drop them. 1914-19 evidence from Schafer, Gaines and Barbee, *Wheat production*, 19, 21.

A6. Additional information on the bushel weight-yield relationship

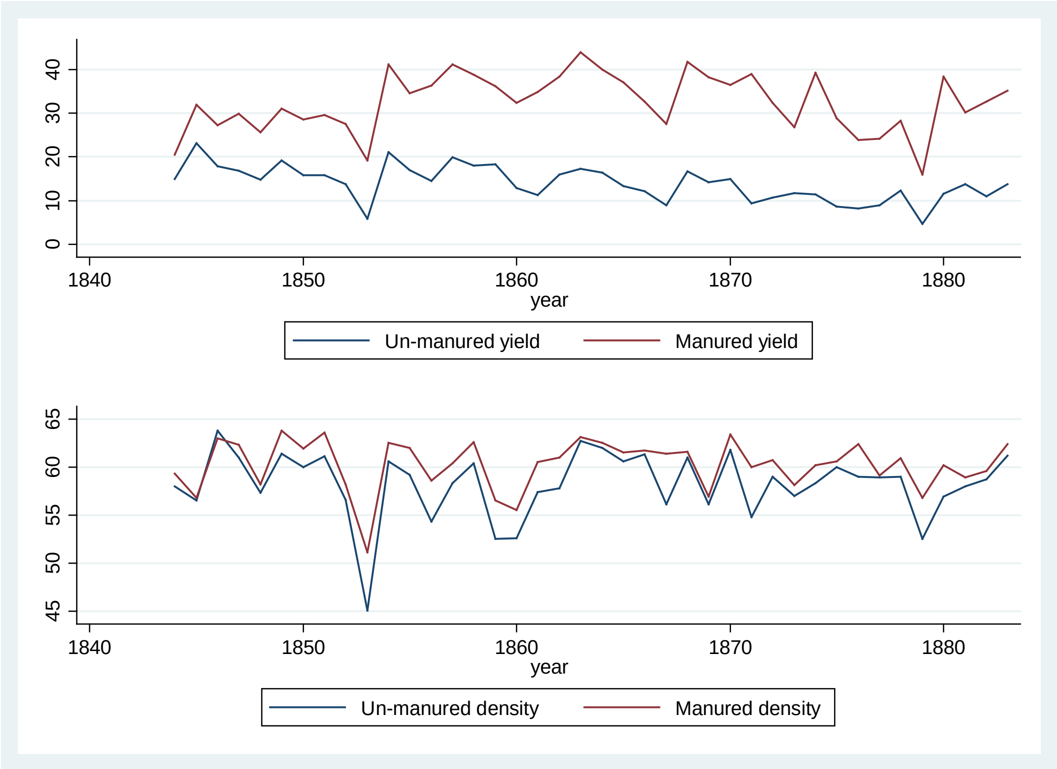
Figure A6.1 graphs data underlying year-on-year yield-bushel weight relationship estimates, drawn from Lawes' and Gilbert's post-1843 Rothamsted perpetual wheat experiments. We use the plot receiving 14 tons/acre/year of dung and the unmanured plot, and the 1871-81 cultivar trials.

Figure A6.1. Scatterplot of the Rothamsted wheat cultivar trial data.



Notes. Scatter plot of Lawes and Gilbert wheat cultivar data set, Rivett's highlighted because Walton (1999) comments it was particularly affected by wireworm in 1879. These data are used for panel estimation (bottom half of table 10).

Figure A6.2. Yields on Rothamstead experimental plots.



Note the unmanured plot yield trending slightly downward (the soil became exhausted as it never received fertilizer); 1853 is problematic, but yields were low beforehand. Wireworm affected 1879 and 1880. Hence we also consider the sub-sample 1854-1878. Main text regressions use the manured plot and both plots pooled: for comparison, unmanured plot regressions are shown below.

Table A6.1 Regression of bushel weight on yield (all in natural logarithms).

	(1) unmanured all	(2) unmanured subsample	(3) manured all	(4) manured subsample	(5) pooled all	(6) pooled subsample
Yield	0.098*	0.028	0.084*	0.037	0.056***	0.037***
	(0.039)	(0.032)	(0.036)	(0.035)	(0.014)	(0.006)
Constant	3.810***	3.993***	3.808***	3.974***	3.911***	3.973***
	(0.105)	(0.079)	(0.128)	(0.126)	(0.046)	(0.022)
N	40	25	40	25	80	50
r^2	0.270	0.026	0.177	0.033	0.261	0.182

A7. Details of the Monte Carlo simulations and additional results

Monte Carlo simulations are based on a cointegrating VAR of the form:

(A.10) $$\Delta \mathbf{p}_t^* = \boldsymbol{\alpha\gamma}\mathbf{p}_{t-1}^* + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \quad \forall t \neq s : \mathrm{cov}\left[\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_s\right] = 0,.$$

where $\mathbf{p}_t^*$ is a vector of log-prices for constant-quality wheat, defined in equation (1), combined with measurement error of the form:

(A.11) $$\mathbf{p}_t^V \equiv \mathbf{p}_t^* + \boldsymbol{\eta}_t .$$

where $\mathbf{p}_t^V$ is observed price. For simplicity, consider the case with just two prices, A and B, and price homogeneity so that $\boldsymbol{\gamma} = \begin{pmatrix} 1 & -1 \end{pmatrix}$. Kuzin and Hassler show that, even with measurement error, $\boldsymbol{\gamma}$ is still estimated consistently and tests for the rank of $\boldsymbol{\alpha\gamma}$ are unaffected so long as sufficiently many extra lags of $\Delta\mathbf{p}$ are included to remove serial correlation. To see why loadings estimators will be inconsistent, substitute (A.11) into (A.10) to obtain:

(A.12) $$\Delta \mathbf{p}_t^V = \boldsymbol{\alpha\gamma}\mathbf{p}_{t-1}^V + \boldsymbol{\mu} + \left\{\boldsymbol{\eta}_t - \left(\mathbf{I} + \boldsymbol{\alpha\gamma}\right)\boldsymbol{\eta}_{t-1} + \boldsymbol{\varepsilon}_t\right\}.$$

Both the Johansen procedure and OLS estimator are inconsistent because the disturbance in this regression – i.e. the term in brackets – is serially correlated and the regression contains a lagged dependent variable.

Remaining assumptions about disturbances and errors in the Monte Carlo are:

(A.13) $$\begin{aligned} &\mathrm{corr}\left[\varepsilon_t^A, \varepsilon_t^B\right] = 0.5; \quad \mathrm{var}\left[\varepsilon_t^A\right] = \mathrm{var}\left[\varepsilon_t^B\right] = \sigma_\varepsilon^2 \\ &\mathrm{E}\left[\boldsymbol{\eta}_t\right] = \mathbf{0}; \quad \mathrm{var}\left[\boldsymbol{\eta}_t\right] = \sigma_\eta^2\mathbf{I}; \quad \mathrm{cov}\left[\boldsymbol{\eta}_t, \boldsymbol{\eta}_{t-1}\right] = \mathbf{0} \end{aligned}$$

Thus there is some correlation between the two disturbances to the true price series (invariably observed in real data) and disturbances have the same variance for both prices. Quality variation is modelled as classical measurement error in both price series. Both disturbances and measurement errors are assumed Normally distributed. We considered a variety of parameter values for equation (A.11). In particular, we considered four measurement error possibilities: no measurement error, then measurement error with a standard deviation of a quarter, half or equal to the disturbance:

(A.14) $$\sigma_{\eta} \in \left\{0, \sigma_{\varepsilon}/4, \sigma_{\varepsilon}/2, \sigma_{\varepsilon}\right\}$$

The “half life” is the time taken for half of a (true) disturbance to die away (i.e. to move halfway back to equilibrium). The relationship of half life to loadings is:

(A.15) $$HL = \frac{\ln\left(0.5\right)}{\ln\left(1 + \boldsymbol{\gamma}\boldsymbol{\alpha}\right)} = \frac{\ln\left(0.5\right)}{\ln\left(1 + \alpha_{1} - \alpha_{2}\right)}$$

where:

(A.16) $$\alpha_{1} \leq 0; \quad \alpha_{2} \geq 0; \quad 0 < \alpha_{1} - \alpha_{2} < 1$$

For any desired half life, we choose parameters satisfying:

(A.17) $$\alpha_{1} - \alpha_{2} = \left(0.5\right)^{1/HL} - 1$$

For example, for a half life of two, the asymmetric and symmetric versions of the model are respectively:

(A.18) $$\boldsymbol{\alpha} = \begin{pmatrix} 0 \\ 1 - \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \frac{1}{2}\left(\frac{1}{\sqrt{2}} - 1\right) \\ \frac{1}{2}\left(1 - \frac{1}{\sqrt{2}}\right) \end{pmatrix}$$

Each Monte Carlo simulation had 20,000 replications. We considered sample sizes of 50, 100 and 200. Figures A7.1 and A7.2A compare sample size effects on dynamic response, graphing median impulse response functions for the error correction: sample size makes little difference in either case.

Figure A7.1. Effect of sample size on impulse response functions: asymmetric loadings

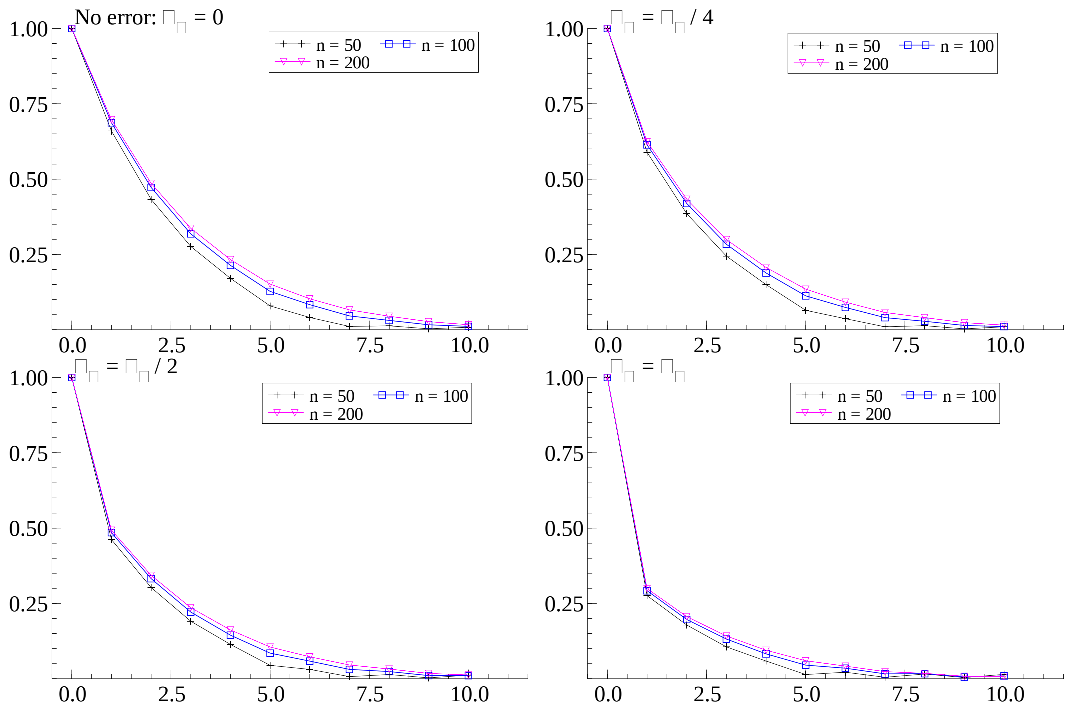
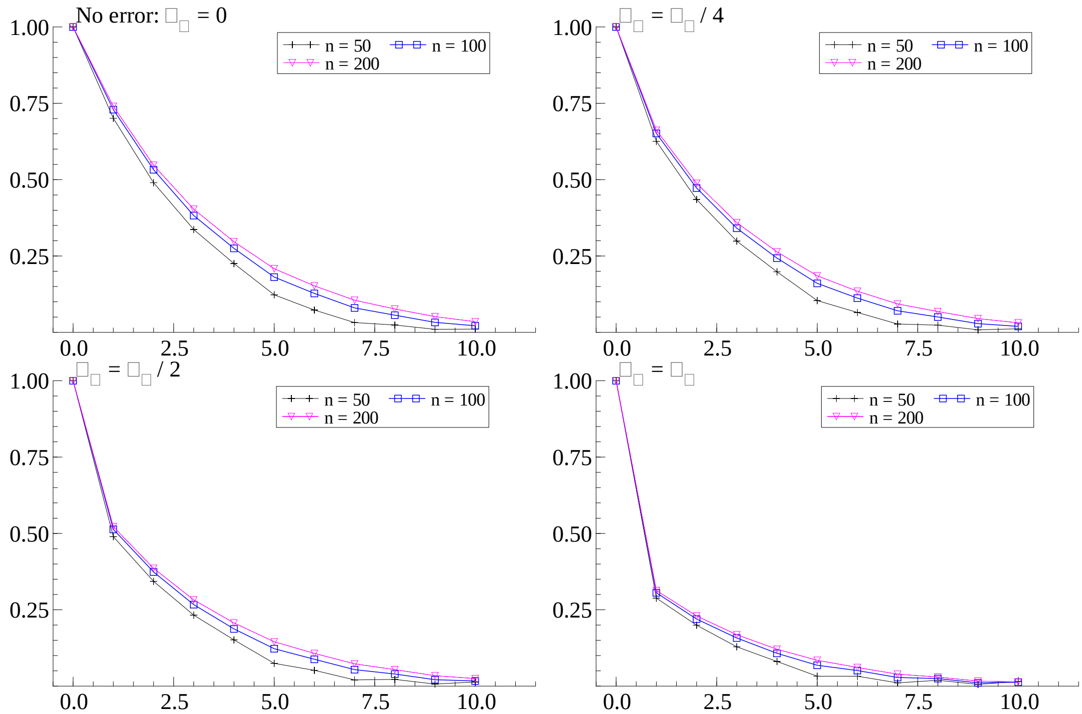


Figure A7.2 Effect of sample size on impulse response functions: symmetric loadings



Deciding what ratio $\sigma_\eta / \sigma_\varepsilon$ to use is more important, since this cannot be observed directly:

any parameter estimates from equation A.13 would be biased, making the ratio $\sigma_\eta / \hat{\sigma}_\varepsilon$ is unreliable. Using our Monte Carlo simulations, we estimate the relationship between $\sigma_\eta / \sigma_\varepsilon$ and σ_η / σ_p . Table 9 reports this ratio for a half life of two and sample size of 200. Table A7.1 provides analogous figures for other sample sizes and both short and long half lives.

Table A7.1 Monte Carlo estimates of σ_η / σ_p

$\sigma_\eta / \sigma_\varepsilon$	0	0.25	0.50	1.00
Sample size = 50				
Half life = 1	0.000	0.231	0.403	0.579
Half life = 10	0.000	0.238	0.413	0.586
Sample size = 100				
Half life = 1	0.000	0.229	0.400	0.574
Half life = 10	0.000	0.236	0.410	0.581
Sample size = 200				
Half life = 1	0.000	0.228	0.399	0.572
Half life = 10	0.000	0.236	0.409	0.579

The final Monte Carlo results here are connected to variance of the parameter estimates. The main text includes information only on the median impulse response. Tables A7.1 to A7.3 summarize variation in impulse response functions (instead of underlying parameter estimates) by reporting the range of the 95% confidence interval. Note this is approximately symmetrically distributed around the median (or mean): one measure of variation is width of the 95% confidence interval.

Table A7.1 reports impulse response for ten periods after a period 0 shock. With no measurement error and a sample size of 200, after one period the impulse response is about a half (not reported) and the range of the 95% confidence interval is 0.32 to 0.66. Width of this range (one period post-shock) is 0.34 to 0.35, regardless of measurement error magnitude. In later time periods the range gets gradually smaller as the shock effect dies away: the magnitudes are fairly similar. With a sample size of 50, the range is about twice as large (unsurprisingly – with the sample size only a quarter of the previous experiment, the range increases by, very approximately, the square root of four).

Effects are more complicated with larger half lives. The fact the half life is larger means there is a slightly larger bias (from the standard results for auto-regressive processes with serial correlation and a lagged dependent variable) and confidence intervals are also wider. However, presence of measurement error biases downwards both half lives and range of estimates.

The effect is not large with a half life of two. Suppose measurement error standard deviation is a quarter of true price disturbance standard deviation. True value of the 95% confidence interval is 0.26 ($n = 200$) or 0.53 ($n = 50$) but measurement error makes it appear 0.23 or 0.48 respectively. If measurement error is larger, the effect for $t = 1$ is to increase the confidence interval, but by $t = 10$ the interval is almost halved. So with measurement error, the precision of the estimates may be overestimated.

Table A7.2 Variation in impulse response functions when true half life equals unity

$\sigma_\eta / \sigma_\varepsilon$		0			0.25			0.50			1.00	
t	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval
Sample size = 200												
0	1	1		1	1		1	1		1	1	
1	0.66	0.32	0.34	0.61	0.27	0.34	0.51	0.17	0.34	0.36	0.00	0.35
2	0.42	0.05	0.37	0.40	0.03	0.37	0.34	-0.01	0.36	0.26	-0.09	0.35
3	0.30	-0.08	0.38	0.29	-0.09	0.37	0.26	-0.11	0.36	0.22	-0.14	0.36
4	0.24	-0.13	0.37	0.23	-0.14	0.37	0.21	-0.14	0.35	0.19	-0.16	0.34
5	0.17	-0.14	0.31	0.16	-0.14	0.30	0.15	-0.14	0.29	0.14	-0.14	0.28
6	0.14	-0.12	0.26	0.13	-0.11	0.24	0.11	-0.09	0.20	0.07	-0.06	0.13
7	0.10	-0.09	0.20	0.09	-0.09	0.18	0.07	-0.07	0.14	0.05	-0.05	0.10
8	0.08	-0.07	0.14	0.07	-0.06	0.13	0.06	-0.05	0.10	0.04	-0.03	0.07
9	0.05	-0.05	0.10	0.05	-0.04	0.09	0.04	-0.04	0.08	0.03	-0.03	0.06
10	0.05	-0.03	0.07	0.04	-0.02	0.07	0.04	-0.01	0.05	0.03	-0.01	0.04
Sample size = 50												
0	1	1		1	1		1	1		1	1	
1	0.84	0.07	0.76	0.78	0.01	0.77	0.70	-0.08	0.78	0.57	-0.25	0.81
2	0.61	-0.21	0.82	0.59	-0.23	0.82	0.54	-0.26	0.80	0.47	-0.34	0.81
3	0.49	-0.33	0.82	0.47	-0.33	0.80	0.45	-0.34	0.79	0.40	-0.36	0.76
4	0.42	-0.37	0.79	0.41	-0.36	0.77	0.39	-0.36	0.74	0.36	-0.36	0.73
5	0.30	-0.34	0.63	0.29	-0.33	0.62	0.27	-0.32	0.59	0.27	-0.32	0.59
6	0.26	-0.28	0.54	0.25	-0.27	0.51	0.23	-0.23	0.46	0.21	-0.19	0.40
7	0.21	-0.24	0.45	0.20	-0.23	0.43	0.18	-0.21	0.39	0.17	-0.18	0.35
8	0.20	-0.19	0.38	0.19	-0.17	0.36	0.17	-0.16	0.33	0.16	-0.13	0.29
9	0.17	-0.17	0.33	0.16	-0.16	0.31	0.14	-0.14	0.28	0.13	-0.12	0.25
10	0.16	-0.12	0.28	0.15	-0.11	0.26	0.14	-0.09	0.23	0.13	-0.07	0.20

Table A7.3 Variation in impulse response functions when true half life equals two

$\sigma_\eta / \sigma_\varepsilon$		0			0.25			0.50			1.00	
t	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval
Sample size = 200												
0	1	1		1	1		1	1		1	1	
1	0.91	0.57	0.34	0.83	0.49	0.34	0.69	0.35	0.35	0.49	0.13	0.36
2	0.76	0.34	0.42	0.69	0.28	0.41	0.58	0.19	0.39	0.42	0.04	0.37
3	0.63	0.17	0.46	0.58	0.14	0.44	0.49	0.08	0.41	0.36	-0.02	0.38
4	0.53	0.06	0.47	0.49	0.04	0.45	0.42	0.00	0.42	0.31	-0.06	0.37
5	0.41	0.00	0.41	0.37	-0.01	0.38	0.30	-0.03	0.33	0.22	-0.06	0.29
6	0.34	-0.05	0.39	0.31	-0.05	0.36	0.26	-0.05	0.30	0.17	-0.04	0.21
7	0.29	-0.08	0.36	0.26	-0.07	0.33	0.21	-0.06	0.27	0.14	-0.04	0.19
8	0.25	-0.08	0.33	0.23	-0.07	0.30	0.18	-0.06	0.24	0.12	-0.04	0.16
9	0.21	-0.09	0.30	0.19	-0.08	0.27	0.15	-0.06	0.21	0.10	-0.03	0.13
10	0.18	-0.08	0.26	0.17	-0.06	0.23	0.13	-0.04	0.18	0.09	-0.01	0.10
Sample size = 50												
0	1	1		1	1		1	1		1	1	
1	1.08	0.31	0.77	1.00	0.23	0.77	0.88	0.09	0.79	0.70	-0.13	0.83
2	0.95	0.02	0.93	0.90	-0.02	0.92	0.78	-0.10	0.88	0.61	-0.23	0.84
3	0.84	-0.17	1.01	0.79	-0.18	0.97	0.69	-0.22	0.91	0.55	-0.28	0.83
4	0.75	-0.26	1.01	0.71	-0.27	0.97	0.61	-0.28	0.89	0.49	-0.31	0.80
5	0.58	-0.30	0.88	0.53	-0.29	0.81	0.43	-0.28	0.70	0.33	-0.28	0.61
6	0.50	-0.30	0.79	0.45	-0.28	0.73	0.39	-0.24	0.63	0.29	-0.19	0.48
7	0.42	-0.31	0.73	0.39	-0.28	0.67	0.33	-0.24	0.57	0.24	-0.19	0.43
8	0.38	-0.28	0.66	0.35	-0.25	0.60	0.29	-0.21	0.50	0.22	-0.15	0.37
9	0.33	-0.27	0.60	0.30	-0.25	0.54	0.25	-0.20	0.45	0.17	-0.15	0.32
10	0.29	-0.24	0.53	0.27	-0.21	0.48	0.23	-0.16	0.39	0.17	-0.09	0.26

Table A7.4 Variation in impulse response functions when true half life equals ten

$\sigma_\eta / \sigma_\varepsilon$		0			0.25			0.50			1.00	
t	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval	97.5%	2.5%	Interval
Sample size = 200												
0	1	1		1	1		1	1		1	1	
1	1.09	0.75	0.34	1.00	0.66	0.34	0.84	0.49	0.35	0.62	0.26	0.36
2	1.08	0.62	0.46	0.98	0.54	0.44	0.83	0.40	0.43	0.60	0.21	0.40
3	1.05	0.51	0.54	0.96	0.44	0.52	0.80	0.33	0.48	0.58	0.16	0.42
4	1.03	0.41	0.61	0.93	0.36	0.57	0.78	0.26	0.51	0.56	0.13	0.43
5	0.96	0.36	0.60	0.86	0.32	0.54	0.69	0.25	0.44	0.47	0.15	0.32
6	0.90	0.30	0.60	0.81	0.27	0.54	0.65	0.20	0.45	0.44	0.12	0.32
7	0.84	0.24	0.60	0.76	0.21	0.55	0.61	0.16	0.45	0.41	0.09	0.33
8	0.80	0.19	0.61	0.72	0.16	0.56	0.58	0.13	0.45	0.40	0.07	0.32
9	0.75	0.14	0.61	0.68	0.12	0.56	0.55	0.09	0.45	0.37	0.05	0.32
10	0.71	0.10	0.61	0.64	0.09	0.55	0.52	0.07	0.45	0.36	0.04	0.31
Sample size = 50												
0	1	1		1	1		1	1		1	1	
1	1.25	0.47	0.77	1.16	0.37	0.79	1.02	0.21	0.81	0.82	-0.03	0.85
2	1.27	0.23	1.04	1.17	0.17	1.00	1.02	0.06	0.96	0.79	-0.11	0.90
3	1.27	0.05	1.22	1.17	0.02	1.15	1.01	-0.06	1.07	0.77	-0.16	0.93
4	1.26	-0.07	1.33	1.16	-0.09	1.26	0.98	-0.14	1.12	0.74	-0.20	0.94
5	1.17	-0.13	1.30	1.06	-0.14	1.19	0.85	-0.14	0.99	0.57	-0.17	0.74
6	1.10	-0.17	1.27	0.99	-0.17	1.16	0.81	-0.15	0.96	0.56	-0.13	0.69
7	1.03	-0.22	1.24	0.92	-0.21	1.12	0.74	-0.18	0.92	0.51	-0.14	0.65
8	0.98	-0.23	1.21	0.88	-0.21	1.09	0.72	-0.17	0.89	0.49	-0.12	0.61
9	0.92	-0.26	1.17	0.82	-0.23	1.05	0.67	-0.19	0.86	0.46	-0.12	0.58
10	0.88	-0.26	1.14	0.79	-0.22	1.02	0.65	-0.16	0.82	0.45	-0.09	0.53

Issued in the series Discussion Papers 2014

2014

01/14 January**, Kurt R. Brekke**, Tor Helge Holmås, and Odd Rune Straume, "Price Regulation and Parallel Imports of Pharmaceuticals".

02/14 January, **Alexander W. Cappelen, Bjørn-Atle Reme, Erik Ø. Sørensen,** and **Bertil Tungodden, "**Leadership and incentives".

03/14 January, **Ingvild Almås, Alexander W. Cappelen, Kjell G. Salvanes, Erik Ø**. **Sørensen**, and **Bertil Tungodden**, "Willingness to Compete: Family Matters".

04/14 February, **Kurt R. Brekke**, Luigi Siciliani, and Odd Runde Straume, "Horizontal Mergers and Product Quality".

05/14 March, **Jan Tore Klovland**, "Challenges for the construction of historical price indices: The case of Norway, 1777-1920".

06/14 March, Johanna Möllerström, **Bjørn-Atle Reme**, and **Erik Ø. Sørensen**, "Luck, Choice and Responsibility".

07/14 March, Andreea Cosnita-Langlais and **Lars Sørgard**, "Enforcement vs Deterrence in Merger Control: Can Remedies Lead to Lower Welfare?"

08/14 March, **Alexander W. Cappelen, Shachar Kariv, Erik Ø. Sørensen**, and **Bertil Tungodden**, «Is There a Development Gap in Rationality?"

09/14 April, **Alexander W. Cappelen**, Ulrik H. Nielsen, **Bertil Tungodden**, Jean-Robert Tyran, and Erik Wengström, "Fairness is intuitive".

10/14 April, **Agnar Sandmo**, "The early history of environmental economics".

11/14 April, **Astrid Kunze**, "Are all of the good men fathers? The effect of having children on earnings".

12/14 April, **Agnar Sandmo**, "The Market in Economics: Behavioural Assumptions and Value Judgments".

13/14 April, **Agnar Sandmo**, "Adam Smith and modern economics".

14/14 April, Hilde Meersman, **Siri Pettersen Strandenes**, and Eddy Van de Voorde, "Port Pricing: Principles, Structure and Models".

15/14 May, **Ola Honningdal Grytten**, "Growth in public finances as tool for control: Norwegian development 1850-1950"

16/14 May, **Hans Jarle Kind**, Tore Nilssen, and **Lars Sørgard**, "Inter-Firm Price Coordination in a Two-Sided Market".

17/14 May, **Stig Tenold**, "Globalisation and maritime labour in Norway after World War II".

18/14 May, **Tunç Durmaz,** "Energy Storage and Renewable Energy"

19/14 May, **Elias Braunfels**, "How do Political and Economic Institutions Affect Each Other?"

20/14 May, Arturo Ormeño and **Krisztina Molnár**, "Using Survey Data of Inflation Expectations in the Estimation of Learning and Rational Expectations Models"

21/14 May, **Kurt R. Brekke**, Luigi Siciliani, and Odd Rune Straume, "Hospital Mergers with Regulated Prices".

22/14 May, **Katrine Holm Reiso**, "The Effect of Welfare Reforms on Benefit Substitution".

23/14 June, Sandra E. Black, Paul J. Devereux, and **Kjell G. Salvanes**, "Does grief transfer across generations? In-utero deaths and child outcomes"

24/14 June, Manudeep Bhuller, Magne Mogstad, and **Kjell G. Salvanes**, «Life Cycle Earnings, Education Premiums and Internal Rates of Return".

25/14 June, **Ragnhild Balsvik**, **Sissel Jensen**, and **Kjell G. Salvanes**, "Made in China, sold in Norway: Local labor market effects of an import shock".

26/14 August, **Kristina Bott, Alexander W. Cappelen, Erik Ø. Sørensen**, and **Bertil Tungodden**, "You've got mail: a randomized field experiment on tax evasion"

27/14 August, **Alexander W. Cappelen, Sebastian Fest, Erik Ø. Sørensen**, and **Bertil Tungodden**, "The freedom to choose undermines the willingness to redistribute."

28/14 August, Marianne Bertrand, Sandra Black, **Sissel Jensen**, and Adriana Lleras-Muney, "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labor Market Outcomes in Norway."

29/14 August, **Astrid Kunze**, "The family gap in career progression".

30/14 September, **Ragnhild Balsvik** and **Morten Sæthre**, "Rent Sharing with Footloose Production. Foreign Ownership and Wages Revisited".

31/14 October, **Nicola D. Coniglio** and Giovanni Pesce, "Climate Variability and International Migration: an empirical analysis"

32/14 November, **Kurt R. Brekke**, Armando J. Garcia Pires, Dirk Schindler, and Guttorm Schjelderup, "Capital Taxation and Imperfect Competition: ACE vs. CBIT"

33/14 November, **Jan I. Haaland** and Anthony J. Venables, "Optimal trade policy with monopolistic competition and heterogeneous firms".

34/14 December, Rolf Aaberge, **Kai Liu**, and Yu Zhu, "Political Uncertainty and Household Savings".

2015

01/15 January**,** Antonio Mele**, Krisztina Molnár,** and Sergio Santoro, "On the perils of stabilizing prices when agents are learning".

02/15 March, **Liam Brunt**, "Weather shocks and English wheat yields, 1690-1871".

03/15 March, **Kjetil Bjorvatn, Alexander W. Cappelen**, Linda Helgesson Sekei, **Erik Ø. Sørensen**, and **Bertil Tungodden**, "Teaching through television: Experimental evidence on entrepreneurship education in Tanzania".

04/15 March**, Kurt R. Brekke**, **Chiara Canta**, Odd Rune Straume, "Reference pricing with endogenous generic entry".

05/15 March, Richard Gilbert and **Eirik Gaard Kristiansen**, "Licensing and Innovation with Imperfect Contract Enforcement".

06/15 March, **Liam Brunt** and Edmund Cannon, "Variations in the price and quality of English grain, 1750-1914: quantitative evidence and empirical implications".

Norges Handelshøyskole

Norwegian School of Economics

NHH
Helleveien 30
NO-5045 Bergen
Norway

Tlf/Tel: +47 55 95 90 00
Faks/Fax: +47 55 95 91 00
nhh.postmottak@nhh.no
www.nhh.no