Structural change out of Agriculture: Labor Push versus Labor Pull

Web Appendix

Francisco Alvarez-Cuadrado   Markus Poschke**

November 2010

D   Data sources

The main sources we rely on are Mitchell (1988, 2003a, 2003b, 2003c) and the Groningen Growth and Development Centre (GGDC) 10-sector and Historical National Accounts databases. They are described in this section. We complement these with additional sources described below if these are reliable and provide longer series.

D.1   Main sources

Mitchell (1988, 2003a, 2003b, 2003c) provides data on the history of sectoral labor allocations in many countries, sometimes going back until 1800. After 1960, Mitchell mainly draws on national statistical yearbooks. For the period up to 1960, the main source is Bairoch and colleagues (1968), who in turn draw on national censuses. While they have made "every possible effort to achieve international and intertemporal comparability" (Mitchell 2003b, p. 144), "frequent changes in criteria and methods used in census

---
**Department of Economics, McGill University, 855 Sherbrooke Street West, Montreal, QC, Canada H3A 2T7. emails: francisco.alvarez-cuadrado@mcgill.ca, markus.poschke@mcgill.ca.

taking [make it] practically impossible to come up with statistics that are perfectly comparable in time and space" (Bairoch 1968). Comparing trends and orders of magnitude, however, is feasible. As a consequence, while the precision of numbers for employment shares in agriculture before 1960 should not be overstated, the main patterns of the data that we present in Section 5 of the paper should be correct. Focusing mainly on stages of the transformation rather than on precise figures for the employment share is also in line with data quality.

For historical relative prices, the Groningen Growth and Development Centre Historical National Accounts database is very useful. Data are available at `http://www.ggdc.net/databases/hna.htm` and described in Smits, Woltjer and Ma (2009). The dataset was created "to bring together the available, but fragmented, data on GDP at the industry level for all major economies and to standardise these series to make a consistent long run international comparison of output and productivity feasible". The database contains series on GDP at current and at constant prices in local currency for the total economy and broad sectors in parts of the 19th and 20th century (about 1870 to 1950 for most countries) using a common industrial classification (2-digit NACE rev. 1.1). All sources are country-specific; detailed references are given below. The GGDC considers the dataset a complement to Maddison's World GDP series and to its 10-sector database and the EUKLEMS data which are both discussed below. The GDP series at current and at constant prices allow imputing sectoral and aggregate price indices. Whereas some of the earlier-generation estimates (essentially the ones on which the series for Germany and for the UK are based) may suffer from some methodological problems mainly due to changes in relative price levels between countries, this is less of a problem for the more recent estimates for other countries. Moreover, given that we use ratios of price levels (p_m/p_a), these problems should be attenuated in our use of the data. Other issues to be borne in mind when using historical data are discussed in Section 3.

Another important source we draw on for post-war labor allocations and relative prices is the GGDC 10-sector database. For Europe, this is an update of van Ark (1996), for Asia it is described in Timmer and de Vries (2007). The data and documentation are available on `http://www.ggdc.net/dseries/10-sector.html`. The database covers 1950 to 2005. It contains information on employment and value added at current and constant prices in ten main sectors of the ISIC rev. 2 classification, one of which is "agriculture, hunting, forestry and fishing". Together, the ten sectors cover the whole economy. We use

the series on employment in agriculture and on total employment, and use the two value added series to impute sectoral and aggregate price indices. Employment is defined as "all persons employed". This includes all paid employees, the self-employed and family workers, which is particularly important for agriculture. For Europe and Japan, the database draws on OECD National Accounts and OECD Labour Force Statistics and on National Accounts of the individual countries. For South Korea, it mainly draws on the Bank of Korea's National Accounts and before that on data from the Economic Planning Bureau. Data thus largely come from sources designed for comparability, use common definitions (employment concept), and a harmonized sectoral classification.[1] Series are thus consistent over time and comparable across countries. Timmer and de Vries (2007) provide a detailed comparison of their data with data from the World Bank World Development Indicators and show that the GGDC data has fewer gaps and a more consistent methodology that suffers less from changes in methodology or survey coverage, which occasionally lead to implausible numbers or jumps in WDI data. This database is preferable to the EUKLEMS database described below (see the entry on Belgium) because it provides a longer time series.

D.2 Sources by country

D.2.1 Employment shares in agriculture

Belgium: 1846-1960: Economically active population by major industrial groups, Agriculture, forestry and fishing and Total, Mitchell (2003*b*, Table B1).

1970-2005: Employment in Agriculture, forestry and fishing and Total Employment, EUKLEMS database, November 2009 release, described in O'Mahony and Timmer (2009) and available at `http://euklems.net`. This database is based on national accounts and labor force survey data. It is similar to the GGDC 10-sector database in focussing on maintaining intertemporal and international comparability. Containing more disaggregate information, the EUKLEMS time series is shorter, starting in 1970 rather than 1960.

Canada: 1881-1971: Population of working age and either gainfully occupied (1881-1941) or labour force (1951-1971), in non-agricultural and agricultural pursuits,

---

[1]While there are minor differences in how countries construct industry-level numbers, these do not tend to change over time and are more of a problem in services.

census years, Historical Statistics of Canada, Table D1-7. Publicly available at `http://www.statcan.gc.ca/cgi-bin/IPS/display?cat_num=11-516-X`.

1976-2006: Total employed, all industries and agriculture respectively, Statistics Canada CANSIM series v2057603, v2057605. Labour force survey estimates.

Finland: 1880-2000: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (2003*b*, Table B1).

France: 1856-1936: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (2003*b*, Table B1).

1950-2005: Employment in Agriculture, Forestry and Fishing and Total Employment, GGDC-10 sector database.

Germany: 1849-1939: Employment by sector from Hoffmann (1965, p. 204), which draws on German Statistical Yearbooks. In the period where they overlap (1882-1939), this series is fairly close to that in Mitchell (2003*b*, Table B1). Fremdling (2008) shows that Hoffmann's 1939 number for the level of employment in agriculture includes (annexed) Austria and Sudetenland, unlike the numbers for previous years. Using data from the *Statistisches Jahrbuch für das Deutsche Reich* (1942, p. 411) to correct this leads to an employment share in agriculture just slightly below that reported by Hoffmann and in line with that reported by Mitchell.

1950-1990: Employment in Agriculture, Forestry and Fishing and Total Employment, GGDC-10 sector database.

Japan: 1872-2000: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (2003*a*, p. 97).

Netherlands 1800-1913: Labour Force, agriculture and total, Smits, Horlings and van Zanden (2000, Table C.1). This monograph summarizes results from the project 'Reconstruction of the National Accounts of the Netherlands', which involved more than 20 members for around a decade. The project used a variety of sources in order to compile a set of consistent historical national accounts. The numbers on the employment share in agriculture are fairly close to the series in Mitchell (2003*b*, Table B1) in the years 1849-1909 where they overlap.

1920-1947: Labor force in agriculture and fisheries and total labor force, van Zanden (1998, Table 6.5): In his monograph on Dutch economic history in the 20th century, van Zanden extends the previous series, drawing on work by authors also involved in the National Accounts project. Where they overlap (1909-1947), this series features a slightly lower level and a smoother decline of the employment share in agriculture than Mitchell (2003b, Table B1).

1950-2005: Employment in Agriculture, Forestry and Fishing and Total Employment, GGDC 10-sector database. This series lies between the levels of data cited in van Zanden (1998, Table 9.4) and that in Mitchell (2003b, Table B1). Mitchell's series features a rather abrupt decline in L^A between 1947 and 1971.

The series we use thus put part of the decline in L^A at an earlier date than the alternative series by Mitchell, with an L^A that is about 3 percentage points lower around 1910-20 and then declines more slowly. However, the difference occurs within a "stage" of our empirical specification and therefore does not affect our results.

South Korea: 1918-1943: Share of households in agriculture, personal communication from Nak Nyeon Kim of Dongguk University, South Korea. The series improves on the series used in Cha and Kim (2006). For this period and place, the share of households (rather than employment) is considered a more reliable measure.

1955-2005: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (2003a, p. 97).

Spain: 1860-2001: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (2003b, Table B1).

Sweden: Number of employed, including self-employed, in agriculture and ancillaries and in the total economy, Edvinsson (2005), who draws on Statistics Sweden. Data available at `www.historicalstatistics.org`. Compared to the series for Sweden by Mitchell, this series features a higher employment share in agriculture until 1890, and therefore faster structural change in the 19th century. The stage dummies used in Section 5 are robust to these differences.

UK: 1801-1910: Share of the labor force in agriculture, Clark (2002, Table 3). The series in fact extends back to 1500. From 1801 on, it draws on population censuses, so we use this more reliable part of the series.

1920-30: Economically active population by major industrial groups, Agriculture, Forestry and Fishing and Total, Mitchell (1988).

1948-2005: Employment in Agriculture, Forestry and Fishing and Total Employment, GGDC 10-sector database.

USA: 1800-1900: Labor force in agriculture and total labor force, Weiss (1992, 1993): The series relies on censuses and improves on earlier series in the classification of some groups of laborers.

1909-1947: Farm employment and total employment (14 years old and over), U.S. Department of Commerce (1975, series D5 and D6). The series actually starts in 1900 and reports an agricultural employment share of 41% for that year, compared to only 36% according to Weiss (1993). This difference is due to a different age cutoff (10 vs 14 years). Following Dennis and İşcan (2009), we use Weiss's numbers until 1900 and the Department of Commerce numbers starting in 1909. (From 1910 on, the employment share in agriculture for 10- vs 14-year olds is virtually identical (Tostlebe 1957).)

1948-2000: Agricultural employment and total employment (16 years old and over), Bureau of Labor Statistics.

The series for the 20th century is close to that in Carter, Gartner, Haines, Olmstead, Sutch and Wright (2006, Table Ba652-669), which however has a lower frequency.

D.2.2   Prices

Belgium: 1836-1953: Price indices for agricultural output and for total output, imputed from the GGDC Historical National Accounts Database, which for Belgium mainly draws on Horlings (1996).

1970-2005: price indices for gross value added in agriculture and the total economy, EUKLEMS database, November 2009 release (O'Mahony and Timmer 2009, also see above).

Canada: 1936-1960: Price indices for agricultural output and for total output, imputed from Statistics Canada CANSIM series v501034, v501035, v11752 and v11753 (GDP at factor cost by industry, current and constant prices, Agriculture and Total industries).

Finland: 1860-2000: Price indices for agricultural output and for total output, imputed from the GGDC Historical National Accounts Database, which for Finland draws on Hjerppe (1989, 1996).

France: Price indices for agricultural output and for total output. 1815-1938: imputed from the GGDC Historical National Accounts Database, which for France draws on Toutain (1987).

1946-1995: Price indices for value added (total and agriculture, fisheries and forestry) from Villa (1993), with data available on the CEPII website at `http://www.cepii.fr/francgraph/bdd/villa/mode.htm`.

Germany: 1852-1913: Price indices for agricultural output, imputed from value added in agriculture, forestry and fishing in current and constant prices (Hoffmann 1965, p. 331, 333) and for net national product (p. 598).

1968-1990: Price indices for value added in agriculture and the total economy, imputed from the GGDC-10 sector database.

Japan: Price indices for agricultural output and for total output, for 1885-1940 imputed from the GGDC Historical National Accounts Database, which for Japan draws on Ohkawa, Takamatsu and Yamamoto (1974), and for 1950-2000 from the GGDC 10-sector database.

Netherlands: Price indices for agricultural output and for total output, for 1808-1939 imputed from the GGDC Historical National Accounts Database, which for the Netherlands draws on various sources, in particular Smits et al. (2000), and from 1970-2005 from the GGDC 10-sector database.

South Korea: 1913-1940: Price indices for agricultural output and for total output, imputed from total GDP and GDP in agriculture in current and constant prices, Cha and Kim (2006, Tables 1 and 2).

1970-2005: Price indices for agricultural output and for total output, imputed from series on GDP in current prices and its growth rate in chained prices, agricultural share of GDP in current prices and growth of agricultural value added in chained (agricultural) prices, available from the Korean Statistical Information service (`www.kosis.kr`).

Spain: Price indices for agricultural output and for total output, for 1850-1950 imputed from the GGDC Historical National Accounts Database, which for Spain draws on Prados de la Escosura (2003), and for 1970-2001 from the GGDC 10-sector database.

Sweden: 1800-2000: Price indices for agricultural output and for total output, imputed from Swedish Historical National Accounts, described in Krantz and Schön (2007) and available at `http://www.ehl.lu.se/database/LU-MADD/National%20Accounts/default.htm`.

UK: Price indices for agricultural output and for total output, for 1861-1938 imputed from Mitchell (1988, Tables 16.2 and 16.8), which draw on Deane and Cole (1962) and Feinstein (1972), and for 1960-2005 imputed from the GGDC 10-sector database.

USA: There are no sectoral value added price series for the U.S., so we rely on wholesale prices. (In Maddison's (1995, Appendix B) words, "we have the paradox that the USA is one of the few countries where the construction of historical accounts by industry of origin has been neglected, though the statistical basis for such estimates is better than elsewhere.")

1800-1890: Wholesale price indices for all commodities and for farm products (Hanes 2006b). These series result from a project by George F. Warren and Frank A. Pearson (1933) to create wholesale price indices for the 19th century that would "correspond with" the Bureau of Labor Statistics wholesale price indices, which go back to 1890. (The series up to 1797 is due to Herman M. Stoker and uses the same methodology.) Sources are mainly newspapers and government reports, and prices refer to New York City. (Series for other cities are also available but are less reliable.) Both this series and the early-19th century BLS series cover few finished goods and no services. However, according to Dennis and İşcan (2009), other price indices with more comprehensive coverage (but more sparse observations) exhibit similar trends.

1890-2000: Wholesale and producer price indexes for all commodities and for farm products (Hanes 2006a). This series links two series from the BLS Handbook of Labor Statistics. From 1947 on, the series consists in the BLS producer price indices by commodity group. This is linked to the wholesale price indices collected by the BLS before that.

References

Bairoch, P. (1968), *The Working Population and its Structure*, Editions de l'Institut de Sociologie, Université Libre de Bruxelles, Brussels.

Carter, S. B., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R. and Wright, G., eds (2006), *Historical Statistics of the United States, Millennial Edition On Line*, Cambridge University Press, Cambridge.

Cha, M. S. and Kim, N. N. (2006), 'Korea's first industrial revolution, 1911-40', *Naksungdae Institute of Economic Research Working Paper Series* **2006-3**.

Clark, G. (2002), 'The Agricultural Revolution and the Industrial Revolution: England, 1500-1912', *mimeo* .

Deane, P. and Cole, W. A. (1962), *British Economic Growth 1688-1959: Trends and Structure*, Cambridge University Press, Cambridge, UK.

Dennis, B. N. and İşcan, T. B. (2009), 'Engel versus Baumol: Accounting for Structural Change using Two Centuries of U.S. Data', *Explorations in Economic History* **46**(2), 186–202.

Edvinsson, R. (2005), *Growth, Accumulation, Crisis: With New Macroeconomic Data for Sweden*, Almqvist & Wiksell International, Stockholm.

Feinstein, C. H. (1972), *National Income, Expenditure and Output of the United Kingdom 1855-1965*, Cambridge University Press, Cambridge, UK.

Fremdling, R. (2008), '(Re)Construction Site of German Historical National Accounts: German Agricultural Employment, Production and Labour Productivity', *Groningen Growth and Development Centre Research Memorandum* **GD-94c**.

Hanes, C. (2006*a*), Wholesale and producer price indexes, by commodity group: 1890-1997 [Bureau of Labor Statistics], *in* S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch and G. Wright, eds, 'Historical Statistics of the United States, Millennial Edition On Line', Cambridge University Press, Cambridge. Table Cc66-83.

Hanes, C. (2006b), Wholesale price indexes, by commodity group: 1749-1890 [Warren and Pearson], *in* S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch and G. Wright, eds, 'Historical Statistics of the United States, Millennial Edition On Line', Cambridge University Press, Cambridge. Table C113-124.

Hjerppe, R. (1989), *The Finnish Economy 1860-1985: Growth and Structural Change*, Bank of Finland, Helsinki.

Hjerppe, R. (1996), 'Finland's Historical National Accounts 1860-1994: Calculation Methods and Statistical Tables', *Publications of Finnish History* **24**.

Hoffmann, W. G. (1965), *Das Wachstum der Deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts*, Springer, Berlin.

Horlings, E. (1996), 'Estimates of Value Added in the Belgian Service Sector 1830-1953', *Catholic University of Leuven Research Memorandum* .

Krantz, O. and Schön, L. (2007), *Swedish historical national accounts 1800-2000*, Almqvist & Wiksell International, Stockholm.

Maddison, A. (1995), *Monitoring the World Economy: 1820-1992*, OECD Development Centre, Paris.

Mitchell, B. R. (1988), *British Historical Statistics*, Cambridge University Press, Cambridge.

Mitchell, B. R. (2003a), *International Historical Statistics: Africa, Asia & Oceania, 1750-2000*, 4th edn, Palgrave Macmillan, New York.

Mitchell, B. R. (2003b), *International Historical Statistics: Europe, 1750-2000*, 5th edn, Palgrave Macmillan, New York.

Mitchell, B. R. (2003c), *International Historical Statistics: The Americas, 1750-2000*, 5th edn, Palgrave Macmillan, New York.

Ohkawa, K., Takamatsu, N. and Yamamoto, Y. (1974), Vol. 1 National Income, *in* K. Ohkawa, M. Shinohara and M. Umemura, eds, 'Estimates of Long-Term Economic Statistics of Japan Since 1868', Tokyo Keizai Shinposha, Tokyo.

O'Mahony, M. and Timmer, M. P. (2009), 'Output, Input and Productivity Measures at the Industry Level: the EU KLEMS Database', *The Economic Journal* **119**(538), F374–F403.

Prados de la Escosura, L. (2003), *El Progreso Económico de España 1850-2000*, Fundación BBVA, Bilbao.

Smits, J.-P., Horlings, E. and van Zanden, J. L. (2000), 'Dutch GNP and its Components 1800-1913', *Groningen Growth and Development Centre Research Monograph* **5**.

Smits, J.-P., Woltjer, P. and Ma, D. (2009), 'A Dataset on Comparative Historical National Accounts, ca. 1870-1950: A Time-Series Perspective', *Groningen Growth and Development Centre Research Memorandum* **107**.

Statistisches Jahrbuch für das Deutsche Reich (1942), Statistisches Reichsamt, Berlin.

Timmer, M. P. and de Vries, G. J. (2007), 'A Cross-Country Database For Sectoral Employment And Productivity In Asia And Latin America, 1950-2005', *Groningen Growth and Development Centre Research Memorandum* **GD-98**.

Tostlebe, A. S. (1957), *Capital in Agriculture: Its Formation and Financing since 1870*, Princeton University Press for the NBER, Princeton.

Toutain, J.-C. (1987), *Le produit intérieur brut de la France de 1789 à 1982*, Économies et Société, Grenoble.

U.S. Department of Commerce (1975), *Historical Statistics of the United States: from Colonial Times to the Present (1970)*, Bureau of the Census, U.S. Government Printing Office, Washington, D.C.

van Ark, B. (1996), Sectoral Growth Accounting and Structural Change in Post-War Europe, *in* B. van Ark and N. Crafts, eds, 'Quantitative Aspects of Post-War European Economic Growth', CEPR/Cambridge University Press, Cambridge, pp. 84–164.

van Zanden, J. L. (1998), *The Economic History of the Netherlands 1914-1995: A Small Open Economy in the 'Long' Twentieth Century*, Routledge, London, New York.

Villa, P. (1993), *Une analyse macro-économique de la France au XXème siècle*, CNRS, Paris.

Weiss, T. (1992), U.S. Labor Force Estimates and Economic Growth, 1800-1860, *in* R. E. Gallman and J. J. Wallis, eds, 'American Economic Growth and Standards of Living before the Civil War', University of Chicago Press, Chicago.

Weiss, T. (1993), 'Long-Term Changes in U.S. Agricultural Output Per Worker, 1800-1900', *Economic History Review* **46**, 324–341.