

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27 **Under-Enumeration, Inconsistency and Bias in the US Manufacturing Census 1860-1880:**

28
29
30 **Case Studies from the American Manufacturing Belt**

31
32
33
34
35
36 **Richard G. Healey**

37
38
39 **Department of Geography, University of Portsmouth**

40
41
42 **Buckingham Building, Lion Terrace, Portsmouth, PO1 3HE, UK**

43
44
45 **Email: richard.healey@port.ac.uk**

Abstract

A new typology of major types of enumeration and tabulation errors found in the manuscript schedules and the published reports of the US Manufacturing Census 1860-1880 is proposed. This is based on a review of historical and contemporary assessments of census accuracy and additional new findings. Detailed cases studies of different manufacturing sectors in the rapidly growing city of Cleveland, OH, including railroad shops, transportation-related manufacturing and oil refining are undertaken using non-census sources, as worked examples of census reporting problems, and comparisons are provided with cities in surrounding states. Significant under-enumeration, bias towards small companies and inconsistent reporting, both within and between census years emerges as the norm rather than the worst case scenario. The manufacturing census is found to be unsuitable for use in quantitative studies without extensive additional corroborating evidence of accuracy.

Keywords: U.S. Manufacturing Census; Under-Enumeration; American Manufacturing Belt; Railroads; Statistical Analysis

Acknowledgements

The author would like to thank the referees for their constructive and insightful comments on the manuscript.

Introduction

Data derived from the 19th century manufacturing censuses of the United States have become a mainstay of quantitative studies of industrial development in historical geography and urban and economic history (e.g. Atack 1987; Klein and Crafts 2011; Meyer 1989, 1990; Niemi 1974). Within the period, the census years 1860-1880 have proved to be especially significant, because the published census volumes contain classified tables of manufacturing activity by county across all states (U. S. Census Office 1865, 1872, 1883a). More limited or selective tabulations (e.g. by major cities) can be found in earlier and later census publications (e.g. U. S. Census Office 1854, 1895a). The potential value of the published reports for research purposes has been further enhanced by the careful work of Bateman, Weiss and Atack (n.d.) to construct a large sample of industrial concerns from the original manuscript manufacturing census schedules for the years 1850-1880. The digital files for the sample records are publicly accessible, in conjunction with the necessary code tables.

The relative accessibility of these datasets, in paper or digital form has therefore added to their allure as the basis for a variety of quantitative studies ranging from the use of locational quotients to sophisticated panel regression models (e.g. Kim 1995, 1999; Meyer 1983). While these studies have provided a wide range of interesting insights into processes of industrial development, the impressive rigour of the analytical approaches adopted has not, in general, been matched by an equally rigorous assessment of the quality of data on which the results are based. Formerly, when data processing limitations restricted most studies to perhaps an individual industry or a particular urban area, such considerations tended to be masked by more general questions about the wider applicability of findings from specific case studies (e.g. Walkowitz 1978). With the growing availability of large digital census and administrative boundary datasets for multiple time periods, such as those from NHGIS, however, the possibilities for much broader scale analyses have expanded considerably

1
2
3 (Minnesota Population Center 2004). Under these circumstances, it is essential to have a
4
5 detailed understanding of potential variability in data quality over time and across space.
6
7

8
9 It may be argued that such problems of data quality are implicit in all historical data
10
11 sources and are well understood by users of such data. So they do not present an impediment
12
13 to the use of census datasets, which have the advantage that they are based on systematic and
14
15 largely comprehensive enumeration of manufacturing establishments, and hence are the best
16
17 available sources of information on 19th century economic development. This would be a
18
19 defensible position, if indeed all the various sources of possible error had been examined and
20
21 found to be limited in size and broadly consistent across the different censuses.
22

23
24 Unfortunately, no such checks appear to have been made and the literature on manufacturing
25
26 census data errors is very small, with the partial exception of Walsh (1970), and Atack and
27
28 Bateman (1999). Inwood (1995) provides a valuable comparison, in terms of the Canadian
29
30 manufacturing census. This limited literature is surprising, given, for example, the lengthy
31
32 treatise in the published 1870 manufacturing census on the shortcomings in different aspects
33
34 of the enumeration (U. S. Census Office 1872, 371f.), which should be required reading for
35
36 all students of 19th century US industrialisation.
37
38

39
40 This study begins to remedy the lack of checking for problems and inconsistencies by
41
42 providing a more systematic examination of a wide range of potential sources of error in the
43
44 manufacturing census. This will include discussion of the specific problems identified in the
45
46 1870 Census Office report. An attempt will then be made to assess the numerical
47
48 significance and wider implications of some of these classes of error. This will be undertaken
49
50 using several detailed case studies from Ohio and adjacent states in the American
51
52 Manufacturing Belt (AMB) (De Geer 1927), since these and similar areas have attracted
53
54 particular attention from researchers who have made extensive use of the manufacturing
55
56 census. These case studies will utilise non-census data sources for comparative purposes. It
57
58
59
60

1
2
3 is also recognised that there may be other classes of reporting problem specific to different
4 parts of the country, such as the ante-bellum South or the pioneer West, but space restrictions
5 necessarily limit the discussion to the AMB, because of the detailed nature of the analysis
6 undertaken.
7
8
9
10

11
12
13 There are two main dimensions to the investigation of enumeration errors and
14 reporting problems in the manufacturing census. The first is to distinguish within-census
15 from between-census problems, while remembering that even within a single census,
16 enumeration was conducted at a large number of geographical locations approximately
17 simultaneously, by different groups of census employees, of varying enthusiasm, knowledge
18 and experience. The second is to separate problems arising during the original ‘door-to-door’
19 enumeration, recorded on the manuscript schedules, from those attributable to the subsequent
20 clerical processing of these schedules, to generate the categorised summary tabulations
21 published by the Census Office for states, counties and in many cases, major cities. These
22 two main problem dimensions are summarised in Tables 1 and 2 respectively and the detailed
23 content of these tables will be discussed further below.
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 Although the existing literature provides some *qualitative* appreciation of the
39 problems faced by enumerators of manufacturing establishments, the possible *quantitative*
40 implications are somewhat less clear. Also, there is very little in the way of surviving
41 documentation about how the manuscript schedules were processed into the published
42 tabulations, and what steps were or were not taken to ensure consistency in reporting both
43 within and between censuses. Therefore possible problems in the later stages of the process
44 may have to be identified or inferred either from an analysis of the printed tabulations
45 themselves, or comparison of these tabulations with other non-census data sources. Given
46 the geographical scale of the census data collection operation, over a restricted time period
47 for each census, in the absence of modern-day information management tools, the scope for
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 many different types of error to enter into the numerous stages of the process between the
4
5 enumerator's visit and the printed tabulation is very easy to envisage.
6
7

8 **Within-Census Problems**

9

10 **The Enumeration Process**

11

12
13
14 Table 1 summarises the main types of both enumeration and tabulation related
15
16 problems that can be identified. Many, but not all of the possible enumeration issues were
17
18 originally detailed by Superintendent Francis Walker (U. S. Census Office 1872, 371f.), as
19
20 can be seen from the first column in the table. A first group of problems derived from the
21
22 fact that the Manufacturing Census was considered the poor relation of the Population Census
23
24 and the resources of time and money devoted to it were deemed inadequate to the
25
26 requirements of the task. In particular, the payment for each firm enumerated had not
27
28 increased since 1850 and was set at a mere 15 cents. This was less than an hour's pay for a
29
30 manual labourer at the time, hardly a sum likely to encourage accurate and comprehensive
31
32 prosecution of the required duties. This was confirmed by subsequent Census Office
33
34 investigations which found 'not a few' cases of unacceptable performance, where under-
35
36 enumeration of firms exceeded 50% and the level of negligence overall was considered great
37
38 enough to 'seriously affect the results of the census' (U. S. Census Office 1872, 372).
39
40
41
42
43

44 One aspect of under-enumeration that is not properly addressed by Walker is whether
45
46 there were other types of more general bias affecting multiple sectors or perhaps urban more
47
48 than rural areas? For example, Stuart made specific reference to the refusal or failure of
49
50 some of the largest firms in the ship-building, sugar refining, paper milling and other sectors
51
52 to furnish information for the 1890 census, though this was believed to be no worse a
53
54 problem than in 1880. This suggests the possible existence of a systematic company size
55
56 bias, although no individual firms were named (Stuart 1898, 624). If enumeration was not
57
58
59
60

1
2
3 consistent in the level of coverage across sectors, locations and size classes, and the extent of
4 this inconsistency could affect data availability for 50% or more of manufacturing concerns,
5
6
7 judging by Walker's figures, then the implications for use of the original manufacturing
8
9
10 census schedules are potentially very significant. It is therefore surprising that these
11
12 questions are not really considered, either by Walsh (1970) or Attack and Bateman (1999). In
13
14 the latter case, assiduous attention is given to explaining the methodology for creating the
15
16 samples from the extant manuscript schedules, but the potential statistical effects of varying
17
18 undercounts or biases in the contents of the schedules themselves is not examined, though
19
20 any biases will necessarily be reflected, or even magnified, in the samples, however
21
22 comprehensive they may appear to be. For example, if there is systematic undercounting of
23
24 larger firms in specific sectors, then analyses relating to the existence of scale economies in
25
26 those sectors, based on sample data, could easily produce results of unknown validity, since
27
28 the actual form of the relationship between costs and outputs, whether linear or nonlinear,
29
30 cannot be reliably determined from a biased subset of the firms involved.
31
32
33
34

35 A second group of enumeration problems derived not from negligence, but from
36
37 difficulties in handling more complex forms of company organisation, where there might be
38
39 multiple plant locations and/or a geographical separation between the Head Office and the
40
41 site of manufacturing activity. This raised not only the possibility of non-enumeration, where
42
43 enumerators in multiple locations all assumed their colleagues elsewhere were reporting on
44
45 the entire concern, but also double-counting, if all enumerators involved recorded
46
47 information for the same firm on multiple schedules. Walker mentions the first of these
48
49 reporting scenarios, but not the second and clearly other partial variants can also be
50
51 envisaged. Another specifically geographical implication of such difficulties could be the
52
53 inflation of employment/capital invested figures for one county, by aggregate head office
54
55 reporting of operations, some of which were located in other counties. An excellent example
56
57
58
59
60

1
2
3 of the complexities that could arise in these circumstances is provided by the Pittsburgh and
4
5 Lake Angeline Iron Company, which had its head office in Cleveland, though its iron ore
6
7 mines were in Marquette County, Michigan (Swineford 1876, 206). It is reported as a single
8
9 concern in the 1870 manuscript manufacturing census schedule for Cuyahoga County, OH
10
11 with 175 employees and a capital of \$386,000. However, it does not appear in the printed
12
13 manufacturing census tabulation. Instead, it was transferred to the mining tabulation, which
14
15 appears elsewhere in the printed 1870 census volumes (U.S. Census Office 1872, 783), to
16
17 give the false information that iron ore mining was being carried on in Cuyahoga county,
18
19 which was not the case. However, the company's mines in Michigan *also* appear in the
20
21 manuscript schedules for that state, but as two separate entries, totalling 180 employees and a
22
23 capital of \$300,000. As in the Ohio case, the entries were transferred to the mining tabulation,
24
25 but as two 'establishments' not one (all the schedule entries for iron ore mining and the
26
27 tabulation can be successfully reconciled in this case to demonstrate this point
28
29 unambiguously). This example indicates that there is not necessarily a correspondence
30
31 between the manufacturing schedules and the printed tabulations, double counting with
32
33 locational implications could easily arise and within the double counting, the reporting itself
34
35 could be inconsistent.
36
37
38
39
40

41
42 It should also be noted that the instructions to enumerators were not being followed in
43
44 such cases. These instructions stated clearly, that 'when large manufacturing corporations are
45
46 part in one subdivision and part in another, they should be reported in that subdivision where
47
48 the office is' (U.S. Census Office 1870b, 20). Adherence to this instruction would
49
50 necessarily have reduced the geographical accuracy of reporting even if double-counting was
51
52 avoided.
53
54

55
56 The third group of problems relate to inaccurate or inconsistent reporting of the
57
58 information required by the census, pertaining to the value of outputs and the levels of capital
59
60

1
2
3 and labour inputs. The \$500 reporting threshold was designed to capture virtually all
4
5 establishments that occupied, at a minimum, ‘the entire labor of one able-bodied artisan ... at
6
7 the present prices of labor’ (U.S. Census Office 1870b, 20). Unfortunately, the threshold
8
9 was not adjusted for the aftermath of Civil War inflation. Since outside anthracite mine
10
11 labour, for example, was remunerated at about 8 cents an hour in 1860, but 20 cents in 1870
12
13 (DLWRR Coal Department, 1860, 1870), much smaller concerns could potentially have been
14
15 included at the latter time point, thereby significantly changing the population of firms
16
17 subject to enumeration. Also, since most businesses were still very small in the Civil War
18
19 Era, a limited degree of ‘downward ‘massaging’ of the output figures by individual artisans
20
21 could have had a large effect on the numbers of concerns actually enumerated.
22
23
24
25

26 In the absence of any consistent accounting standards across industrial sectors, not to
27
28 mention variations in the treatment of issues such as asset depreciation, the reporting and
29
30 interpretation of figures for capital invested by firms was necessarily fraught with difficulty.
31
32 The Walker Report also highlighted a number of more specific problems in this regard,
33
34 including major opposition from firms to respond to the question at all, lack of knowledge by
35
36 most small proprietors of the actual figure, and for corporations, failure to distinguish
37
38 between the value of issued share capital at par and current market value of shares, when the
39
40 figures could easily differ by ‘50 or 75 percent’ (U. S. Census Office 1872, 381). Branch
41
42 plants could also report capital differently than head offices, as seen in an earlier example.
43
44 While the enumerators may have encountered opposition from firms, based on inspection of a
45
46 wide variety of schedules from different states, this does not appear to have resulted in
47
48 frequent blank columns of capital data. Instead, many of the figures appear rounded to
49
50 convenient, if not very accurate totals, especially at the lower end of the range, but also at the
51
52 higher end, suggesting that a good deal of rough estimation took place to encourage the
53
54 enumerators on their way. A detailed check on the Bateman-Weiss-Atack 1870 sample
55
56
57
58
59
60

1
2
3 confirmed this tendency. Out of nearly 10,000 firms, more than 27% had a listed capital
4
5 corresponding to one of the \$500 increments between \$500 and \$2000, while larger \$10K
6
7 incremental values from \$10K up to \$100K had many more occurrences than intervening
8
9 values. Unfortunately, it would be extremely difficult to undertake any kind of systematic
10
11 independent checking of these capital figures, e.g. for individual cities or states.

12
13 Occasionally, state bureaux of industrial statistics collected relevant data using their own
14
15 reporting methods, but these had widely varying response rates by firms, and questions about
16
17 capital employed were asked in some years, but not others, or only for selected industrial
18
19 sectors (see for example, the Annual Report of the Secretary of Internal Affairs for 1877,
20
21 533-5, and 1881, 70-75). Therefore, it is difficult to disagree strongly with the final
22
23 assessment of the Walker Report that statistics of capital invested were ‘entirely
24
25 untrustworthy and delusive’ (U.S. Census Office 1872, 381). At best, they are broadly
26
27 indicative in counties where enumeration was assiduously conducted; at worst they may be
28
29 grossly inaccurate and misleading.
30
31
32
33

34
35 The Walker Report also reminded users of the statistics on wages that they only
36
37 represented a fraction of actual labor inputs, because they did not include the inputs of
38
39 artisans, who paid themselves from the profits of the business, though this point was not
40
41 made clear in the printed instructions to enumerators. In some industrial sectors, this issue
42
43 was very significant, as there were more artisans than ‘hired hands’ at work (U.S. Census
44
45 Office 1872, 381).
46
47

48 49 **The Tabulation Process**

50
51
52 While the literature from the 1890s onwards reveals a certain reluctance either to
53
54 challenge the representativeness of the manuscript schedules or to test their veracity or
55
56 comprehensiveness against non-census sources of data, certain aspects of the interface
57
58
59
60

1
2
3 between the manuscript schedules and the published county level reports have received
4 significant and justified critical attention, as can be seen from Table 1. Walsh has identified
5 'gross errors and deficiencies' in processing the data in the schedules into the published
6 county level reports in the Wisconsin 1850 and 1860 census. These included omission of
7 meat-packing in most counties in 1850 and more than one hundred significant disparities
8 between the manuscript schedules and printed tabulations in 1860 (Walsh 1970). In a similar
9 vein, Atack and Bateman have identified evidence of arbitrary massaging of the figures by
10 Washington bureaucrats, errors in county tallies in Florida and Oregon and typographical
11 errors in the California tabulations. As a result they express scepticism about any research
12 findings based solely on the printed county level reports (Atack and Bateman 1999, 186).
13
14
15
16
17
18
19
20
21
22
23
24
25

26 In addition to these problems, which may be specific to individual states or counties,
27 there are more generic issues concerned with the overall methodology used for classifying
28 individual firms into broader categories for tabulation purposes. These have not been widely
29 discussed in the literature. The initial problem is that there is no surviving documentation
30 about how the classification was undertaken, i.e. there is no equivalent for tabulation work of
31 the 'instructions to enumerators'. This provides considerable scope for inconsistency
32 between individual county or state summaries, not least because of the wide range of possible
33 firm types and the fact that many firms produced a variety of products. For example, how
34 should a firm that builds both bridges and railroad cars be classified, or one that produces car
35 wheels and general castings? Questions such as these affect the ability of researchers to
36 match tabulated categories to original entries in the manufacturing schedules, especially if the
37 number of firms in the category is quite large. This in turn can cause difficulties for attempts
38 to determine whether all the firms enumerated have been included in the tabulations.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 Omissions could be inadvertent or potentially deliberate, especially in the case of large firms,
55
56
57
58
59
60

1
2
3 as will be seen in an example in a later section. Either way, the accuracy of specific
4
5 tabulations could be significantly affected.
6
7

8 There are two further aspects of the problem of matching schedules to tabulations.
9
10 The first is that several states lack surviving schedules for some or all counties, mainly in the
11
12 period 1850-70. For example, a substantial proportion of the Ohio counties, covering the
13
14 cities of Cincinnati, Cleveland and Columbus, are missing in 1860 and several counties again
15
16 in 1880. Secondly, from 1880 onwards a new practice of using special agents with industry
17
18 expertise, in addition to general enumerators, was used, mainly in larger cities, to collect data
19
20 on a number of industries (Wright 1900, 63). However, the manner of dealing with the
21
22 special schedules is somewhat confusing. In 1900, for example, the special schedules
23
24 pertaining to eight industries, including flour and grist mills, brickyards and tanneries, were
25
26 simply given to enumerators, but for 24 other industries, e.g. iron and steel, special agents
27
28 were placed in charge of the returns, which were handled by mail, and special industry
29
30 specific reports were later published (U.S. Census Office 1900, 51-3). A similar, if simpler,
31
32 approach was adopted in 1880. It involved both special *schedules* collected by enumerators
33
34 and special *returns* collected by special agents, though the process is not well documented.
35
36 None of the special returns have survived, according to the National Archives (Delle Donne,
37
38 1973 referenced in Atack and Bateman 1999). However, a large number of the 1880 special
39
40 schedules can be found, interleaved with the standard manufacturing schedules for counties
41
42 in New York State, Ohio and Pennsylvania, for example (U.S. Census Office 1880). The
43
44 printed heading for the standard schedules lists ten industries to be entered on special
45
46 schedules instead. However, the special schedules themselves include additional categories
47
48 for meat packing and leather tanning. Although one of the ten special industries on the
49
50 standard schedule heading is listed as 'coal mining', on the actual special schedule in
51
52 question this has become 'small coal mines' and evidence of completed schedules in known
53
54
55
56
57
58
59
60

1
2
3 large-scale coal mining areas, indicates it was not, in fact, used to report coal mining proper.
4
5 A further important issue is that iron and steel works were excluded from the 1880 standard
6
7 schedules. There was no documented special schedule for this sector, yet a substantial
8
9 quantity of information was still collected by special agents and published, *both* as part of the
10
11 standard county by county manufacturing summaries and in one of the special census reports
12
13 (U.S. Census Office 1888). Direct evidence of this exclusion process can be found in
14
15 examples such as West Middlesex Township, Mercer Count, PA, where entries mistakenly
16
17 made on the standard schedule for two blast furnaces and a rolling mill were marked ‘omit’ in
18
19 later marginal annotations. Similarly, in New Ringgold, Schuylkill County, PA, extensive
20
21 details of the blast furnace leased from the Philadelphia and Reading Coal and Iron Company
22
23 were scored out and marked ‘taken by Special Agt’, in large letters (U.S. Census Office
24
25 1880).
26
27
28
29

30
31 In relation to the schedules that now survive, these findings in relation to the mining
32
33 and iron and steel industries (in 1880 only) are somewhat unwelcome. On the positive side,
34
35 county level printed summary data for these sectors have come down to us, in the standard
36
37 manufacturing report, or special reports, or both (U.S. Census Office 1883a; 1886; 1888), but
38
39 set against this, there are no surviving records of the individual company entries that were
40
41 used to create the county level summaries, so the latter cannot be checked using census
42
43 derived information. As the originators of the dataset indicate, this also means that the 1880
44
45 manufacturing census sample (Atack and Bateman 1999, Bateman et al., n.d.) does not
46
47 contain these two sectors and in this respect is less representative than the published county
48
49 tabulations, a point they do not make and which arguably runs counter to their contention,
50
51 noted earlier, about the lesser reliability of the tabulations compared to data derived directly
52
53 from the manuscript schedules. The lack of these two sectors is of considerable numerical
54
55
56
57
58
59
60

1
2
3 significance, since in 1880 coal mining employed 171,464 men and boys and iron and steel
4
5 140,978 hands (U.S. Census Office 1886, xxviii; 1888, 9).
6
7

8 A final issue that needs to be addressed is the unambiguous evidence of serious
9
10 omissions of specific standard categories from individual county tabulations. Many examples,
11
12 which are also important for the later discussion of the transportation sector, can be found in
13
14 the 1870 census, e.g. in relation to a sub-category of machinery, entitled 'railroad repairing'.
15
16 In Crawford County (OH), this 'railroad-repairing' sub-category does not appear, so none of
17
18 the substantial numbers of railroad workers in different shops there, including those of the
19
20 Pittsburgh, Fort Wayne and Chicago Railway in Crestline, can be identified, even though the
21
22 latter railroad has an entry in the manuscript schedules (U.S. Census Office, 1870a; U.S.
23
24 Census Office, 1872, 712-718; Perrin, Battle and Goodspeed, 1881, 500). The same problem
25
26 arises in Marquette County (MI), Allegheny County (PA) (Pittsburgh) and in Wayne County
27
28 (MI), where none of the Detroit railroad shop activities are identifiable, when it is known that
29
30 there were 231 employees at the shops of the Michigan Central RR in the city (Michigan
31
32 Central Railroad Annual Report for 1869/70, 44). It should be noted that these examples
33
34 were the first few checked, not the result of an exhaustive search, so there will doubtless be
35
36 many more cases yet to be uncovered. In some instances, the railroad-related information
37
38 may have been incorporated in other categories, in others, omitted altogether. However, it is
39
40 very difficult, if not impossible to determine the precise status of individual cases, even
41
42 assuming the original schedules survive. What remains clear is that the 1870 tabulations
43
44 have misrepresented railroad-related manufacturing activity to a significant degree.
45
46
47
48
49
50
51
52
53

54 **Between-Census Problems**

55 **The Enumeration Process**

1
2
3 Major types of problems under this heading are listed in the left-hand columns in
4
5 Table 2. The first group of issues relate to the changing nature of industry over time. In the
6
7 latter part of the 19th century, average firm size increased progressively; see, for example,
8
9 Hirsch's study of industrialisation in Newark (Hirsch 1978, 25). In simple monetary value
10
11 terms, this trend was reinforced between 1860 and 1870 by the interaction between inflation
12
13 and the \$500 enumeration threshold discussed earlier. However, the deflation of the late
14
15 1870s (Fels 1959) and the removal of the threshold by 1880 produced an effect in the reverse
16
17 direction. Secondly, firms not only increased in size, but also in the complexity of their
18
19 organization (Chandler 1962, 1977). For example, there is the well-known shift in
20
21 Pennsylvania from scattered small charcoal furnaces during the 1850s to large integrated iron
22
23 and steel works, which operated both multiple blast furnaces on one or more sites and distant
24
25 banks of coke ovens in the Connellsville region (Warren 1973).
26
27
28
29

30 While the nature of industry itself was changing, very little attention has been directed
31
32 to examination of the changing definitions of 'manufacturing' used by the Census Office.
33
34 One key point is clear, the construction of avenues of transportation, namely roads, railroads
35
36 and canals was excluded from manufacturing, as was the 'grading of grounds' (U.S. Census
37
38 Office, 1872, 376). However, the status of workers engaged in the construction of new
39
40 manufacturing plants is less apparent. Another contentious category was activities related to
41
42 repair and maintenance. Stuart notes that in the course of preparing the final publications of
43
44 the 1890 census, comparisons were made of industrial categories reported in 1870 and 1880
45
46 and it was found that certain types of repair work were included in 1870 and 1890, but not in
47
48 1880. Incidentally, the same applied to the millinery, dressmaking and bottling categories
49
50 (Stuart 1898, 623).
51
52
53
54

55 Other evidence of variations in reporting practice can readily be identified. One good
56
57 example is coal mining, which was part of the general manufacturing enumeration in the
58
59
60

1
2
3 years 1850-1870. In 1860, it appears as part of the county manufacturing tabulations, but in
4
5 1870 mining statistics were reported separately by state and county (U.S. Census Office
6
7 1872). In 1880 only the special report on mining survives (U.S. Census Office 1886), not the
8
9 manuscript schedules, apart from isolated sheets. This prevents a consistent sampling
10
11 strategy for mining schedules being implemented across all four census years.
12
13

14
15 Another more complex example, which necessarily overlaps with the first two types
16
17 of tabulation problem on the right-hand side of Table 2, is the case of railroad shops and the
18
19 manufacture of locomotives and rolling stock. A problem is immediately apparent because
20
21 these shops often engaged in the latter types of manufacture, in addition to the day-to-day
22
23 work of repairs and maintenance, which raised difficulties in terms of categorisation.
24
25 Elaborating beyond Stuart's earlier quoted remarks, railroad shops/repairing were not
26
27 enumerated in 1860, whereas in 1870 they were supposed to have been enumerated and
28
29 tabulated as a sub-heading of machinery. Unfortunately, previous examples have shown that
30
31 this reporting requirement was not consistently applied across different states and counties.
32
33 Manufacture of freight and passenger cars appears as a separate heading, but locomotive
34
35 manufacture cannot be identified as such. In 1880, the only relevant category tabulated is
36
37 'cars, railroad, street, and repairs', which does not match the 1870 definition. Railroad shops
38
39 are neither enumerated nor tabulated. Lessons had been learned by 1890, however, and two
40
41 clearly defined categories are tabulated, albeit by states, not counties. These are 'cars and
42
43 general shop construction and repairs by steam railroad companies' and 'cars, steam railroad,
44
45 not including operations of railroad companies'. The manufacture of locomotives cannot be
46
47 separately distinguished. The same categories are utilised in 1900, but additionally there is a
48
49 brief special report on locomotive manufacture. From 1880 onwards, these categories are
50
51 also tabulated, where relevant, for 100, 165 and 209 principal cities respectively (U.S. Census
52
53 Office 1883a; 1895a; 1902). It should also be noted that both the 1880 and 1890 censuses
54
55
56
57
58
59
60

1
2
3 produced special reports on transportation. These volumes provide extremely valuable data
4 about individual railroad companies and railroad expansion within very broad geographical
5 sub-divisions of the USA, but they do not have statistical breakdowns by state or county that
6 would make them comparable to the rest of the manufacturing census (U.S. Census Office
7 1883b; 1895b).
8
9

10
11
12
13
14
15 A second group of important questions about inter-censal comparability were
16 originally raised by North in papers published in the 1890s. In the first (North 1890), he
17 questioned whether successive manufacturing censuses might lack comparability because
18 they were taken at different stages in the cyclical fluctuation of the economy, though he did
19 not pursue the point in any depth. Answers to this question may depend on the timescale
20 considered. Examination of the Ayres index of business fluctuations (Price 1933) suggests
21 that in the short-term view, i.e. a few months on either side of the census month, the 1860 and
22 1870 censuses were at broadly similar points on the rising limb of modest and short-lived
23 periods of business expansion. In contrast, the 1880 census was on the falling limb of a short
24 period of pull-back from an earlier brief expansion. On a longer timescale, 1860 represented
25 the final phase of return to overall growth following the 1857 Panic, while 1870 came at the
26 end of a period of mildly fluctuating business fortunes that had persisted since 1867. The
27 1880 case was different again, as it represented the first phase of the lengthy railroad boom
28 that followed on from the deep depression of 1873-8. It is interesting to note that none of the
29 census years above coincided with major recessions and the statistical good fortune continued
30 in 1890. Thus, despite undoubted differences in capacity utilisation between the census years,
31 it was not the case that one census represented a period of major expansion while another
32 reflected widespread idling of capacity, firm bankruptcies or closure of works, as would have
33 been the case had the census been taken, for example, in the year 1877.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 In the later papers, North noted the strange absence of systematic reporting of
4
5 miscellaneous manufacturing expenses in censuses prior to 1890, which would affect any
6
7 calculations of profit margins made from the data, and the lack of comparable standards of
8
9 measurement for units of machinery capacity between different years (North 1899).
10

11
12 Prior to 1870, firms were asked only about numbers of employees. In 1870, the
13
14 census also recorded the number of full-time-equivalent months of operation. By 1880, the
15
16 scope of employment related questions was significantly expanded to encompass not only the
17
18 maximum number of employees, but also the length of the working day in summer and
19
20 winter, and the extent of full- and part-time operation. Prior to 1880 it is therefore difficult to
21
22 be sure that firms were reporting labour inputs any more consistently than they were capital
23
24 inputs. Available evidence indicates that the working day declined progressively in length
25
26 from 11.5 hours in the 1830s to 10.0 hours by the 1880s (Atack and Bateman 1992). Nearly
27
28 40% of establishments worked for less than the full 12 months in 1880, and the average
29
30 extent of downtime among this group of firms was 3.3-3.5 months (Atack et al. 2002, 797).
31
32 Simple employee counts do not therefore imply comparable labor inputs between different
33
34 firms and it is also not possible to determine how comparable the figures are for censuses pre-
35
36 and post-1880.
37
38
39
40
41

42 **The Tabulation Process**

43
44
45 To an important extent, this heading reflects many of the key difficulties of inter-
46
47 censal comparison in the manufacturing census. Problems of lack of compliance with
48
49 tabulation procedures within a census become compounded when those procedures
50
51 themselves (in terms of the tabulation categories used) vary between censuses. Several
52
53 examples of the interpretative difficulties these issues raise have already been given. To this
54
55 must be added the possibility of variations in small size bias between censuses, i.e. large
56
57
58
59
60

1
2
3 firms may be omitted from the relevant tabulation in one census, even if they were
4
5 enumerated, while they are included in an earlier or later census tabulation. If the data were
6
7 to be taken at face value, especially if other firms in the sector were small, this would
8
9 obviously lead researchers to incorrect conclusions about employment/investment variability.
10
11 To the problem of temporal inconsistency must be added the dimension of geographical
12
13 variability, which has major implications for comparative mapping work. Specific counties
14
15 where tabulation procedures were not followed in one census may comply more rigorously in
16
17 the next, but without detailed domain knowledge or other data sources from outside the
18
19 census, it is impossible to determine where this has happened and which years and industrial
20
21 sectors are affected.
22
23

24
25
26 Lastly, there remains the traditional census analysis problem of changing
27
28 administrative boundaries, which has attracted increased attention since the advent of
29
30 historical GIS methods (Gregory and Ell 2007). While the extraordinary efforts of the Atlas
31
32 of Historical County Boundaries Project at the Newberry Library have documented all these
33
34 changes (Long and Tuck Sinko 2010), their effects still have to be incorporated appropriately
35
36 into analytical studies. Many of the changes are minor, but others, such as the splitting off of
37
38 Lackawanna County from Luzerne County, PA in 1878 had major effects, transferring the
39
40 rapidly expanding industrial and mining city of Scranton into the former, while leaving
41
42 Wilkes-Barre in the latter. Since larger manufacturing firms tended to be concentrated in and
43
44 around urban areas, simple data adjustment methods, such as areal interpolation are less
45
46 satisfactory than in the case of the population census. For sampling based studies, the
47
48 problem has been addressed by only using counties where no boundary changes took place
49
50 during the period of interest (Atack, Haines and Margo 2011, 166). However, for map-based
51
52 studies this is not appropriate and questions remain about the possible introduction of hidden
53
54 biases, e.g. if rapidly growing regions are more likely to experience boundary changes than
55
56
57
58
59
60

1
2
3 those where population is static or declining. GIS point-in-polygon methods could
4
5 theoretically resolve boundary change problems, but for this to work several items of
6
7 information would be required. These would include the exact locations of all establishments
8
9 likely to be affected by boundary changes, whether the establishments were actually
10
11 enumerated and which category they were classified into in the relevant published tabulation.
12
13 Obtaining such information would be a major undertaking, if it were indeed feasible, since in
14
15 many cases involving branch plants enumerated with their head offices, the location of the
16
17 branch plants may not have been given in the census entry.
18
19
20
21
22
23

24 **Wider Implications of Inter-Censal Enumeration and Tabulation/Classification**

25 **Problems**

26
27
28
29
30 The widespread absence of consistency in the treatment of different manufacturing
31
32 sectors in the decades prior to 1890 means that the census by itself cannot provide a reliable
33
34 means either of tracking the growth of specific sectors over time in specific locations or
35
36 comparing one sector to another within and between censuses. As has been demonstrated,
37
38 the problems are particularly acute for the transportation sector, though not confined to it.
39
40
41

42
43 There are two main reasons why researchers need to be concerned about these
44
45 findings. The first is that a number of previous studies have made exactly the kinds of inter-
46
47 censal comparisons whose validity has now been seriously questioned. The second is more
48
49 specific to transportation, in that a number of studies have argued about the relative
50
51 importance of the railroad sector in stimulating other kinds of industrial growth (Cootner
52
53 1963; Fogel 1964; Jenks 1944), so it is important that the evidence base is as accurate as
54
55 possible.
56
57
58
59
60

1
2
3 The problem of inappropriate inter-censal comparisons applies, for example, to the
4 figures reported by Niemi (1974) in his classic and widely referenced study. It also has a
5 major impact on Meyer's (1990) study of cities in the AMB. He compares location quotients
6 for SIC manufacturing categories, based on the county level manufacturing data for 1860 and
7 1880 (Meyer 1990, table 4, 743). The transportation sector, which is one of his SIC
8 categories, forms an important part of his argument about the role of exchange mechanisms
9 in city growth, but in neither of these years does the census contain meaningful figures for
10 this sector. Further to this, since railroad shops were substantial employers, often among the
11 largest in urban areas (Wilson 2005), their complete omission from the 1860 and 1880 census
12 and partial omission from the 1870 dataset means that the digital samples from the
13 enumeration schedules necessarily contain a specific bias away from some important large
14 concerns, which could affect the findings of analyses, such as those reported by Attack,
15 Haines and Margo (2011).
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 Given the potentially large size of concerns involved in transportation-related
33 manufacture, regardless of whether they were railroad shops or independent manufacturing
34 firms, and the lack of accurate census data about them, it is essential to investigate ways of
35 improving the estimates of the scale of these operations using non-census sources. To
36 demonstrate some possible ways in which this can be achieved, a case study of Cleveland,
37 Ohio will now be undertaken, as it was a rapidly growing city during the period (Robison
38 1887; Rose 1950), with a large railroad sector. To avoid over-concentration on railroads per
39 se, some additional investigation of transportation-related industry and oil refining
40 developments will also be undertaken, to see whether the kinds of census inadequacies
41 already identified can be found in these sectors also. The wider significance of these
42 findings in relation to the development of the AMB is examined elsewhere (Healey 2014).
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Derivation of New Employment Estimates for Selected Manufacturing Sectors in**
4
5 **Cleveland 1860-1880**
6
7

8 *The Railroad Sector in 1860-1: Shop Men and Total Railroad Employees*
9

10
11 Prior to 1860 the city of Cleveland in Cuyahoga County on the shore of Lake Erie was
12
13 served by five major railroads. Western connections were provided by the Cleveland and
14
15 Toledo RR (CTRR), while those to the south-west were provided by the Cleveland,
16
17 Columbus and Cincinnati RR (CCCRR). From the east and south-east, traffic entered the city
18
19 via the Cleveland, Painesville and Ashtabula RR (CPARR), the Cleveland and Pittsburgh
20
21 (CPRR) and the Cleveland and Mahoning (CMRR) railroads (Cleveland Board of Trade for
22
23 1865, 5-6). The CTRR and the CPARR were merged into the Lake Shore and Michigan
24
25 Southern Railway (LSMSRR) in 1869 (LSMSRR Annual Report for 1870, 5-6). A map of
26
27 these railroads can be found in Healey (2014). The Cuyahoga County manuscript
28
29 manufacturing schedules are among the missing Ohio counties for 1860 and the printed
30
31 tabulation omits railroad shops, so non-census sources such as city directories, newspaper
32
33 records and State Government Reports have to be used instead, if railroad employment is to
34
35 be estimated for this date. Newspaper reports usually contain evidence of actual visits made
36
37 by reporters to the shops in question, rather than being based on second-hand information
38
39 (see individual references below). State reports, either from the office of the Commissioner
40
41 of Statistics or later the Commissioner of Railroads, indicate from the outset, both the
42
43 establishment of a large, state-wide network of nearly 800 correspondents with local
44
45 knowledge, by the Statistics Bureau, and the use of annual circulars for data collection. The
46
47 latter were followed up with persistence by government officials, in the event of non-
48
49 submission or unsatisfactory reporting (Ohio, Commissioner of Statistics Report for 1858,
50
51 31; for 1859, 3; Ohio, Commissioner of Railroads Report for 1869, 7-8). It is therefore
52
53
54
55
56
57
58
59
60

1
2
3 considered reasonable to use data derived from these sources, as a substitute for data not
4
5 reported by the census.
6
7

8
9 The first task is to establish which railroads had shops in the city by this time, before
10 estimating likely employment levels. As early as March 1857, a newspaper report stated that
11 the CPARR was employing 200 hands in their substantial Cleveland roundhouse and shops
12 complex. Although much lower figures of 60-70 men were reported the following year in the
13 aftermath of the 1857 Banking Panic, employment had rebounded to near 1857 levels by
14 1862 (Annals of Cleveland 1857, 272; 1858, 448; Healey 2014). A further 1858 report noted
15 that the newly established CMRR Shops already had a complement of 50 men, at a time
16 when the company had only recently begun full-scale operations and all the rolling stock was
17 of recent manufacture, so the requirement for repairs was limited (Annals of Cleveland 1858,
18 449). The CCCRR was the largest of the railroads entering Cleveland during the 1850s and
19 had a total workforce of 775 by 1858, compared to 577 for the CPARR and 625 for the
20 CTRR, though a significant proportion of these employees would have been trackmen or
21 workers at depots along the lines, who would have lived outside the city (Ohio,
22 Commissioner of Statistics, Second Annual Report for 1858, 75). There is ample evidence
23 that the CCCRR progressively developed substantial shop complexes over time, but
24 unfortunately there are no precise employment figures for these Cleveland shops prior to
25 1870. The first documented mention of an 'engine shop' for the CCCRR was in 1850, and by
26 the beginning of 1854 a large car shop had been erected and a 'heavy force' was engaged in
27 manufacturing cars as rapidly as they were required, although rolling stock manufacture did
28 not extend to locomotives at this time (Annals of Cleveland, 1850, 333; CCCRR Annual
29 Report for 1853, 4). Late in 1855, the railroad commenced grading for another engine house
30 and machine shop near its terminal facilities on Lake Erie (Annals of Cleveland, 1855, 487).
31 In 1857, it purchased the locomotive works of the independent Cuyahoga Steam Furnace
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Company on a site in central Cleveland adjacent to the river, so it could manufacture and
4
5 repair its own engines, though there was initially a lack of demand for additional motive
6
7 power in the depressed conditions following the 1857 Banking Panic (Annals of Cleveland,
8
9 1862, 582; Joblin, 1869, 398; CCCRR Annual Reports for the years 1857, 1859, 1860). A
10
11 further phase in the development of the company's facilities took place in the years 1864-66.
12
13 This involved expansion on land adjacent to the locomotive shops and in the lakeside
14
15 terminal area. Substantial new brick freight and roundhouses were constructed in 1864, in
16
17 addition to machine and repair shops, and the car shops were extended (Annals of Cleveland,
18
19 1864, 233; CCCRR Annual Report for 1864, 14).
20
21
22
23

24 In contrast to the CCCRR, CTRR shop men were unlikely to have been resident in
25
26 Cleveland in 1860, because the company closed its west Cleveland shops in 1857 and moved
27
28 all repair work about 40 miles west to its Norwalk shops (CTRR Annual Report for 1856/7,
29
30 24). Map evidence from 1858 (Hopkins 1858) clearly shows just to the east of the CPARR
31
32 shops the substantial roundhouse and repair shop of the CPRR, established in the early 1850s.
33
34 These shops employed 155 men by 1861 (Healey 2014). By 1868 the overall workforce (not
35
36 just shop-men) of the CPRR had grown substantially to 1,137 employees, almost as large a
37
38 complement as the Cleveland, Columbus, Cincinnati and Indianapolis Road (CCCIRR),
39
40 newly formed by the consolidation of the CCCRR with the Bellefontaine and Indiana RR
41
42 (Ohio, Commissioner of Railroads, Annual Report for 1868, 100, 107).
43
44
45
46

47 The evidence therefore points to a minimum of several hundred railroad shop-men at
48
49 the start of the 1860s, allowing for possible fluctuations with business conditions. A more
50
51 specific, though provisional, estimate of 450-475 shop-men is given by (Healey 2014). This
52
53 initial estimate was used in conjunction with a multiplier of 95.7 other railroad workers to
54
55 every 100 shop-men to infer the presence of 431-455 additional railroad workers, most of
56
57 whom would have lived in Cleveland. This multiplier was based on later data for Cleveland
58
59
60

1
2
3 railroads in the 1880 census (see more detailed explanation below). Adding an allowance for
4
5 CTRR train men (since this railroad did not contribute to the total for shop-men but its tracks
6
7 ran into the city), resulted in an initial estimate of about 900 railroad men in Cleveland at this
8
9 time (Healey 2014), none of whom appear in the Manufacturing Census tabulations.
10

11
12 Since this omission is so large, it is important to examine another independent source,
13
14 in conjunction with the newspaper reports, to determine if the above employment estimate is
15
16 realistic. To this end, analysis of city directory data was undertaken, the closest applicable
17
18 volume being for 1860-1 (Cleveland Directory 1861). No claim of comprehensiveness is
19
20 made for this resource, since other studies in both earlier and later time periods suggest
21
22 directory coverage could vary from about 70 to almost 100% of the corresponding
23
24 enumeration of the white population in the census (Alexander 1982; Goldstein 1954; Knights
25
26 1969). However, for present purposes, it can be utilised to set a lower bound on the size of
27
28 employment in the railroad sector. The first pass through this directory resulted in the
29
30 abstraction of nearly 5,800 industrial and transportation workers out of a total of
31
32 approximately 14,000 directory entries. Simple counts of obviously self-identified railroad
33
34 workers in the resulting list suggested an employment count broadly similar to that given
35
36 above. However, subsequent more detailed analysis of this initial database indicated the need
37
38 to add additional entries to ensure absolutely consistent reporting. Although individuals
39
40 associated with most of the relevant occupations had been correctly incorporated into the
41
42 database, it was found that other broad occupational groupings, such as laborers, and those
43
44 with a range of sub-types, such as agents and clerks, had not been comprehensively included,
45
46 and additional individuals were needed to ensure the comparability of the calculations for
47
48 different occupational sub-groups in the final tabulations presented here. Some 600 further
49
50 entries were therefore added to the final database to make 6,425 records in all. Extremely
51
52
53
54
55
56
57
58
59
60

1
2
3 fine-grained occupational analysis was then conducted in an attempt to refine the initial
4
5 employment estimates and this work is reported here.
6
7

8
9 Previous work by the present author (Healey 2013) has shown that an important issue
10 in the analysis of railroad workers in the census is the difference between railroad specific
11 trainmen (brakemen, conductors etc.), the vast majority of whom identified themselves as
12 railroad men, and generic tradesmen, who overwhelmingly did not. Analysis of the
13
14 Cleveland city directory database reveals exactly the same pattern, as can be seen in Table 3.
15
16 This table is necessarily complex, because the aim is to clarify exactly what can definitely be
17
18 determined about railroad employees and to distinguish it from other estimates about the
19
20 numbers possibly employed in this sector, while also placing bounds on those estimates. At
21
22 the same time, the categories in the table are constrained by the differential levels of
23
24 occupational information provided by individuals responding to questions posed by the
25
26 directory compilers. The guiding principle in table construction is therefore only to
27
28 categorize based on information actually provided in the directory entries, not on inference
29
30 from stated occupations. The key distinction, looking across the table in the horizontal, is
31
32 between individuals who did not identify themselves as railroad workers and those who did
33
34 (either in general or specifically as working for a named railroad). A second level sub-
35
36 division in each of the two main categories is then between those who indicated a work
37
38 location, such as a factory or railroad shop, and those who did not. In the vertical dimension,
39
40 actual or potential trainmen are distinguished from several broad types of generic workers,
41
42 who in turn are separated from tradesmen, as commonly understood, who are also of a
43
44 generic kind. The point about the latter group is that they are potential candidates for
45
46 undeclared railroad shop workers. In addition, there is a final group of ‘other railroad
47
48 workers’ that do not fit into the above more specific occupational categories. While specific
49
50 less common railroad occupations may only appear a few times in the directory, unlike the
51
52
53
54
55
56
57
58
59
60

1
2
3 categories higher up in the table, in sum they form a sizeable group of railroad workers
4
5 overall.
6
7

8 The first key finding can be found in the bottom right-hand corner of the table, which
9
10 indicates that 580 men are unambiguously identifiable as railroad workers, and for the
11
12 majority, the railroad company is also known (see breakdown in Table 4 below). This
13
14 represents the absolute minimum number of such workers in Cleveland at the time. However,
15
16 there are several important pieces of evidence in Table 3 that would suggest the actual total
17
18 was much higher. The first of these is the ‘non-railroad identified potential trainmen’ section
19
20 of rows and columns. These provide a breakdown of 163 individuals working in occupations
21
22 largely associated with railroading, though this is not guaranteed. Thus baggage workers,
23
24 brakemen and conductors, would mostly have been railroad men, but occasionally they could
25
26 be found working for steamship lines on Lake Erie (the ‘vessel worker’ heading), or on street
27
28 railroads. Engineers and firemen are more problematic. Both groups were frequently
29
30 railroaders, but the directory provides clear evidence for the former group, at least, that this
31
32 appellation was also used by workers in factories and on steamboats. Since there is
33
34 insufficient evidence to make close estimates of the real number of railroad men in each
35
36 occupation within this overall grouping, a conservative, but somewhat arbitrary, figure of
37
38 two-thirds of the 140 potential trainmen of unspecified work location (93 men) has been
39
40 taken, as a reasonable indication of the number likely to have been railroad employees.
41
42
43
44
45

46 The non-railroad-specific generic workers present a different problem. The term
47
48 ‘agent’ covered a wide spectrum of related activities, but as a result, the associated directory
49
50 entries were often much more specific as to employment sector, e.g. book agent or insurance
51
52 agent. Hence, a majority of these individuals were definitely not railroad workers. The clerks
53
54 follow a somewhat similar pattern, although a smaller proportion give a specifically non-
55
56 railroad work location. Laborers are different again. Very few give a work location,
57
58
59
60

1
2
3 suggesting a lack of identification with any specific employer, as might be expected when
4 many were doubtless day laborers. Unfortunately, their total numbers are very large, which
5 introduces a significant element of uncertainty into any estimation of their potential
6 contribution to railroad employment. One way of addressing this problem is to find evidence
7 from other railroads of similar size, to serve as a guide to likely employment levels. The
8 Boston and Maine Railroad is a good example. In 1861 it employed 585 men (Boston and
9 Maine Railroad 1861, 36-7), of whom 43 were laborers at the freight and passenger depots at
10 the terminus in Boston. Although Cleveland itself was smaller than Boston, it also provided
11 substantial terminal facilities for five railroads, which on average were each of broadly
12 similar size to the Boston and Maine, both in terms of total earnings and employment, though
13 freight earnings tended to be higher for the Cleveland railroads and passenger earnings lower.
14 Applying the Boston and Maine figure to the Cleveland railroads would give 215 such
15 workers (43 x 5), of whom only 32 are specifically accounted for on the right hand side of
16 Table 3, leaving 183 as a proportion (15%) of the 1218 unspecified laborers on the left hand
17 side of the table. Applying this conservative proportion to the generic workers category as a
18 whole would attribute an estimated 236 of these men (1576 x 0.15) to the railroad sector.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 The category of generic tradesmen raises at least two major issues. The first of these
40 is that of the non-railroad identified workers, only 14.2% give a workplace, and if the slightly
41 special case of the 116 ship carpenters in the 'vessel workers' column is removed, this
42 percentage drops to 6.9%. Hence a large number of skilled tradesmen (1358), in quite
43 specific occupations relevant to railroad work, remain as candidates to be undeclared railroad
44 employees. Secondly, only 45 generic tradesmen identified themselves as railroad workers,
45 and of these, only 22 gave their workplace as the shops. The complete absence of
46 blacksmiths, carpenters and tanners (who worked on passenger car roofs), and the presence of
47 only three recorded shop-based boilermakers and 14 such machinists confirms there is a
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 reporting problem. Earlier conservative estimates, based on a number of newspaper reports,
4 suggested there were 450-475 railroad shop-men, so the remaining 428-453, say 440 in round
5 figures, who are missing (after subtracting the 22) must have made up about a third of the
6 large pool of 1358 unallocated generic tradesmen found in the directory. This is not fanciful,
7 since two of the newspaper reports of the period give employee breakdowns that support the
8 argument about under-reporting. For example, in early 1861 the CPRR shops in Cleveland
9 employed 41 machinists, 21 blacksmiths and 24 carpenters (Cleveland Morning Leader, Feb.
10 19, 1861), while in late 1862, the corresponding figures for the CPARR shops were 34, 15
11 and 40 men (assuming carpenters can be equated to 'wood shop' workers). The CPARR
12 shops alone also employed 15 boilermakers at this time, but there is no indication of a
13 significant contingent of shop laborers (Cleveland Morning Leader, Oct 21, 1862). Other
14 evidence, not reliant on the Cleveland newspapers, confirms that blacksmiths, machinists and
15 carpenters were employed in quantity by railroads of the period. For example, in 1857 the
16 Baltimore and Ohio, though twice the size of all five Cleveland railroads combined, had 138,
17 315 and 268 workers in these three occupations respectively (Healey 2013, 23).

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37 The final tally of all the sub-categories is summarized in Table 5. Double counting
38 has been avoided, because the 'estimated' column largely relates to data on the left hand side
39 of Table 3, while the 'identified' column simply contains the category sub-totals from the
40 right hand side of that table. Several additional points arise from these figures. Firstly, no
41 attempt has been made to extrapolate from known to unidentified depot workers, as there is
42 no extant evidence available in sufficient detail to allow this to be done. That said, it is
43 apparent from the Table 3 that just the known depot men constituted a significant proportion
44 of the railroad workforce in the city. Secondly, there has also been no multiplier applied to
45 the 'other railroad workers' who are outside the depot location column. Some of these were
46 management or specialist workers, such as auditors, and in general the recorded occupations
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 in this group were of low frequency. This prevents any useful generalization about possible
4
5 under-reporting. Finally, despite conservative and category-specific estimation, of a more
6
7 detailed kind than has previously been undertaken for railroad workers using city directory
8
9 sources, the clear finding is that the likely number of railroad workers was 50% greater even
10
11 than the earlier estimate of 900 men (Healey 2014). The latter was based on much simpler
12
13 counts and assumptions than the present analysis. The new estimated total of 1349 men
14
15 (Table 6), not only enhances the argument presented in that earlier study about the
16
17 importance of railroad employment, but in the present context, it makes the omission of any
18
19 reference to railroad work in the 1860 manufacturing census a still more serious problem than
20
21 originally argued above.
22
23
24

25 26 *The Railroad Sector in 1870*

27
28
29 A simpler approach is possible in 1870 because the Cuyahoga County manuscript
30
31 schedules have survived and they contain details of three of the railroad shops in the city,
32
33 though one was not enumerated. Details for the three shops can be matched directly to the
34
35 'railroad-repairing' category in the printed tabulations, so the under-enumeration is reflected
36
37 in both sources, adding a further type of inconsistency to the 1870 census reports. The
38
39 schedules indicate that the CPRR shops employed 160 men, as compared to the 600 at the
40
41 multi-site CCCIRR Shops and 200 at the shops of the Atlantic and Great Western RR (by
42
43 then lessor of the CMRR) (U.S. Census Office 1870a). During this immediate post-bellum
44
45 period, newspaper reports confirm the manufacture of locomotives, rolling stock, and iron
46
47 bridges, in addition to repair work, at the CCCIRR, CPRR and CPARR shops in the city
48
49 (Annals of Cleveland, 1868, 682; 1869, 467-8; Cleveland Board of Trade for 1865, 18).
50
51 Unfortunately, the CPARR shops, which had become part of the operations of the LSMSRR
52
53 by this time, are missing from both the 1870 census schedules and tabulations; so only a
54
55 rough estimate of their employment can be made for this date. However, figures are available
56
57
58
59
60

1
2
3 for 1878. At this time these shops employed 286 people, but the report noted that there had
4
5 been a 'large reduction' in employment levels since 1872 (Ohio, Bureau of Labor Statistics,
6
7 Second Annual Report for 1878, 102). The opening of the new LSMSRR yards in 1874 at
8
9 Collinwood, 7 miles east of the centre of Cleveland, also complicates the position a little.
10
11 These yards were primarily for handling freight trains, and initially contained just a large
12
13 engine house, so a very small number of shop men would have worked there (LSMSRR
14
15 Annual Report for 1873, 13; Encyclopedia of Cleveland History, 1997). However, the main
16
17 repair shops remained within the city of Cleveland as late as 1892 (Cram, Beers and Bennett,
18
19 1892). Thus the 286 employees referred to above can be assumed to be at the main Cleveland
20
21 shops. Given that the number of LSMSRR employees in Ohio had been just over 37% higher
22
23 before the 1873 Panic than after it, when strict economies of operation were introduced
24
25 (LSMSRR Annual Report for 1874, 10; Ohio, Commissioner of Railroads Annual Report for
26
27 1873, 356; for 1880, 922), this suggests the number of shop men working for this railroad
28
29 could have been as high as 400 in the early part of the 1870s. Taken together with the 960
30
31 shop men working for the other three roads, gives an estimated total of about 1360 shop men
32
33 in the city in 1870. This conservative figure compares favourably with the earlier fulsome
34
35 claim by the Board of Trade that in the busy year of 1865 'about 1500 men' were employed
36
37 at railroad shops in the city (Cleveland Board of Trade for 1865, 18).
38
39
40
41
42
43

44 Once a figure for shop men is available it is possible to derive a further estimate of
45
46 total railroad employment in the city, in the following way. For railroads as a whole in 1880,
47
48 the first date for which consistent figures are recorded nationally across all major companies,
49
50 shop men constituted about 21% of total railroad employment, though in the case of the large
51
52 railroads of interest here, the figure ranged from 18-37% (Licht 1983, 34; calculated from
53
54 Table VI in U.S. Census Office 1883b, 261f.). For these same Cleveland railroads,
55
56 comparison of figures for some of the other sub-categories of workers indicates that there
57
58
59
60

1
2
3 were on average 95.7 trainmen and administrative staff for every 100 shop men. Though
4
5 these men were distributed among the main divisional centers, so were the shops, hence it is
6
7 reasonable to use this multiplier on Cleveland shop men in 1870, assuming that the mix of
8
9 employee types had not changed dramatically 1870-1880, to obtain an estimate of Cleveland
10
11 trainmen/administrative staff, i.e. about 1300 men in round figures. Added to the shop men,
12
13 this produces an estimate of 2660 men. The other employment categories of section men
14
15 (track maintenance gangs) and station men would have necessarily been much more spread
16
17 out across the rail network, but directory evidence clearly shows that a significant number of
18
19 the latter were employed at the Union Passenger Depot and the freight depots of the
20
21 individual roads. These workers would be sufficient to take the final estimate of railroad
22
23 employment in Greater Cleveland in 1870 to at least 2750 men (Table 7), dwarfing all other
24
25 employment sectors in the city at the time and representing a doubling of numbers since
26
27 1860/1. The only important footnote to this, and one which is very indicative of the overall
28
29 numbers of railroad men in the city, is that after the opening of the Collinwood yards in 1874,
30
31 some 500 LSMSRR employees (very largely trainmen) are reported to have moved to this
32
33 location in East Cleveland township, though this would have taken place over a period of
34
35 time. Most would have come from central Cleveland itself, but doubtless some also came
36
37 from Painesville (between Cleveland and Ashtabula, see Figure 2 in Healey 2014), as the
38
39 LSMSRR closed their stockyards at this location and moved them to Collinwood
40
41 (Encyclopedia of Cleveland History, 1997; Annals of Cleveland, 1874, 140-1).
42
43
44
45
46
47
48
49
50

51 *The Railroad Sector in 1880*

52
53
54 Unfortunately, the 1880 city directory is too large to process in the same manner as
55
56 the 1861 directory, without the deployment of research resources substantially beyond those
57
58
59
60

1
2
3 currently available. In common with later post-1870 directories in other cities, they also
4
5 become much less informative about places of work, so the information yield progressively
6
7 declines while processing costs rise (cf. Cleveland Directory 1871; 1880). Fortunately
8
9 however, in the case of 1880, the complete individual level population census records are
10
11 available from the North Atlantic Population Project (NAPP) (Minnesota Population Center,
12
13 2008). Previous analysis of occupational codes for railroad-related workers in this dataset
14
15 indicated that the trainmen could be identified quite accurately, but the same problem of
16
17 distinguishing shop men in generic trades arises as in the case of the city directories. Indeed,
18
19 it is a greater problem, since the census rarely includes any workplace, as opposed to
20
21 occupational, information.
22
23

24
25
26 A total of 21 occupational codes used by NAPP (unharmonized OCC variable –
27
28 NAPP-modified HISCO codes) have previously been identified as providing a close
29
30 approximation to railroad-related employment, excluding the generic shop workers (Healey
31
32 2015). These include several codes of very minor numerical importance, which overlap with
33
34 street railroads, such as unspecified conductors and baggage-men. These will inflate the
35
36 figures, but only very slightly. A further 8 codes provide counts of the numbers of individuals
37
38 working in the major generic trades utilized by railroad shops, such as blacksmiths,
39
40 machinists and carpenters, but clearly large numbers of these census individuals worked
41
42 outside the railroad sector, so the railroad proportion has to be estimated, as before. Minor
43
44 generic trades such as painters have been neglected to compensate for the minor over-
45
46 estimation of trainmen just mentioned.
47
48
49

50
51 The method adopted was to calculate the total number of individuals in these different
52
53 groups of codes both for the 54 enumeration districts comprising Cleveland proper, and for
54
55 the 2 enumeration districts comprising East Cleveland, to capture the Collinwood shops
56
57 contingent of workers. A total of 1,874 train service employees are identified in Cleveland
58
59
60

1
2
3 with a further 261 in East Cleveland. The corresponding figures for generic workers were
4
5 6,029, of which one-third were carpenters and joiners, with a mere 70 in East Cleveland.
6
7 Using the previous divisor between trainmen and shop men of 0.957 would give 1,958 shop
8
9 men, which again is similar to the one-third of generic workers being railroad men derived
10
11 above from analysis of the 1861 directory entries. The East Cleveland figures, while much
12
13 smaller, confirm the progressive movement of trainmen to Collinwood, though better-off
14
15 workers may have decided to remain in Cleveland and commute, while the trivial number of
16
17 generic workers (say one-third of 70 individuals) supports the earlier conclusion that only a
18
19 few shop men were employed in connection with the engine house. Thus, a conservative
20
21 estimate of direct railroad employment in Cleveland city in 1880 is at least 3,832 individuals
22
23 (1,874 trainmen + 1,958 shop men) (Table 8), with nearly 300 more in East Cleveland.
24
25
26 Additional individuals working in the depots have also been neglected in these calculations.
27
28 Based on just these Cleveland figures this suggests a less than 50% increase over the decade,
29
30 a much smaller increase than during the 1860s in both absolute and percentage terms.
31
32
33
34

35
36 Depending on whether the documentation for integrated or unharmonized variables is
37
38 consulted on the NAPP website (Minnesota Population Center, 2008), it might be imagined
39
40 that these calculations could be obviated by using the IND50US variable code 506, which
41
42 appears to provide employment counts for the railroad sector directly. However, earlier work
43
44 has shown (Healey 2011) that this is not the case, because IND50US necessarily fails to
45
46 account for the unassigned generic workers, as the required information to do this is not
47
48 present in the census schedules. IND50US therefore overwhelmingly utilizes the same
49
50 information as the OCC codes, as suggested by the documentation for the integrated variable,
51
52 together with a very limited amount of information on work type/location occasionally found
53
54 in the OCCSTRNG occupational transcriptions. These were downloaded separately from the
55
56 main coded datasets and analyzed to determine how the coding was undertaken. The rules
57
58
59
60

1
2
3 used to code OCCSTRNGs are not always easy to discern, sometimes questionable and
4
5 occasionally incorrectly applied (e.g. three railroad callers in East Cleveland, who are
6
7 assigned to a general (not railroad) conductor category (OCC 36090), when callers are not, in
8
9 fact, conductors). For completeness, it can be noted that the 506 code produces 2034 railroad
10
11 workers in Cleveland and 278 in East Cleveland, slightly higher figures than used in the
12
13 previous paragraph, but spreads them across a much wider group of less specific occupational
14
15 codes than the restricted list of OCC codes used above. This makes the identification of the
16
17 appropriate multiplier for shop men much less clear and more difficult to apply. Hence, it was
18
19 found more satisfactory to base the calculations on the OCC code counts directly.
20
21
22

23 **Lessons from the Railroad Sector**

24
25
26
27 Several initial lessons can be drawn from this detailed case study spanning three
28
29 census years. Firstly, the absence of consistent enumeration of railroad shops in both the
30
31 manuscript and published manufacturing census means that no valid conclusions from these
32
33 sources about the growth or otherwise of this part of the transportation
34
35 manufacturing/employment sector can be drawn in Cleveland during this important period
36
37 encompassing the Civil War when the AMB began to emerge (cf. Meyer 1983). Earlier
38
39 comments about census reporting omissions also indicate that this problem affects a number
40
41 of other important manufacturing cities. Secondly, very detailed additional investigation
42
43 using multiple sources separate from the manufacturing census does allow meaningful and
44
45 comparable estimates, at least of employment levels, to be obtained, but the research resource
46
47 implications are high. Thirdly, in this and potentially other cases, the high levels of
48
49 employment identified in the railroad sector, once both shop men and other workers have
50
51 been accounted for, raises important questions about the economic significance of this sector
52
53 in comparison to wider manufacturing activity, which the manufacturing census is unable to
54
55 illuminate.
56
57
58
59
60

Independent Transportation-Related Manufacturing

While it is not the purpose here to investigate linkages between industrial sectors or sub-sectors, it is useful to continue evaluation of the reporting of transportation-related manufacturing to see if the problems identified in the preceding paragraphs were unique to the railroads. In this context the manufacture of rolling stock by independent companies is an appropriate second case study. The 1860 Cuyahoga County tabulation lists only one firm under 'cars', which employed 30 people with a capital of \$25K (U.S. Census Office, 1865, 447) (Table 6). This may have been Morrill and Bowers, although the workforce seems small for a long-established company that produced thousands of cars during the Civil War period and was still employing 70-80 workmen in 1866 (Annals of Cleveland 1866, 643). By 1870, the census tabulation again lists a single concern under 'cars, freight and passenger', but with only 10 employees and \$10K capital. This concern cannot be unambiguously identified by reference to the manuscript schedules, and even if it could, taking the 1860 and 1870 reported figures together suggests a sub-sector on the verge of extinction. However, the 1870 Board of Trade Report conveys a quite different picture. The McNairy and Claflin Manufacturing Company, stated to be the only independent railroad car builder in the city, had produced 1200 freight cars and numerous streetcars during the year and 15,000 lineal feet of railroad and highway bridge superstructure, giving employment to 400 men (Cleveland Board of Trade for 1870, 29). This company does not appear in the manuscript census, which adds weight to earlier concerns about small size bias, since it was one of the largest manufacturing concerns in the city. Thus the 1870 returns have completely misrepresented the scale of car building in the city at that date (Table 7). Further to this, related manufacturing, such as the casting of car wheels by Bowler and Mohr (40 employees) and the making of railroad car roofs by the Winslow Company, the former appearing in the manuscript census, the latter in the Board of Trade Report (1870 Census, Cleveland Board of

1
2
3 Trade for 1870, 37), have not been associated with car manufacture in the printed census
4 report, so these transportation-related manufacturing activities have either been concealed by
5 the classification process for tabulation, or missed in the enumeration. By 1880, four car
6 manufacturers have appeared in the census report, with a combined total employment of 216
7 men, so a comparison with 1870 would give a false impression that a previously moribund
8 sector had somehow been revitalized during the 1870s depression (U.S. Census Office 1883a,
9 322). The printed report can be matched very closely to the manuscript manufacturing
10 census, because employment on car manufacture at the McNairy and Claffin works is
11 helpfully separated from that concerned with bridge manufacture. The population census
12 does not provide a useful check, because in Cleveland the relevant OCC and the IND50US
13 codes are used to cover both car and carriage manufacture and there is further ambiguity in
14 the detailed occupational descriptions (NAPP OCCSTRNG and OCCLABEL) about the
15 location of the work in question. The relevant OCC50US code used (690 – operatives and
16 kindred workers) is too all-encompassing to be of any value here. Almost all the 22
17 individuals describing themselves as ‘car builders’ on the original schedules, are given the
18 closest available OCC code 81300 ‘Coach, Carriage and Wagon Makers’, though it is
19 actually an incorrect designation, and IND50US code 379 ‘Railroad and Misc. Transportation
20 Goods’. However, in Cleveland there are a total of 428 individuals with a 379 code. Of
21 these, only a further 18 were clearly employed in car manufacture, based on their
22 OCCSTRNG values, while 63 were ambiguously recorded with variants of the designation
23 ‘car shop worker’. This could either mean they worked for a car manufacturer, or were
24 employed on car building and/or repairs in a railroad car shop. The remaining individuals
25 were carriage builders or axle works employees. *These sub-counts within the code could only
26 be determined by inspection of the individual transcribed occupational descriptions, thereby
27 rendering the codes themselves of very limited value for present purposes. Thus only 40*

1
2
3 individuals (22+18) can be unambiguously attributed to the independent (i.e. non railroad
4
5 company) car manufacturing sector (Table 8), based on the population census, together with
6
7 an unknown proportion of the other 63 car shop workers recorded. This further indicates that
8
9 50-80% of car manufacturing workers must have declared only the generic nature of their
10
11 work, not its specific location.
12

13
14
15 In Detroit, by way of comparison, the figures for 1870 and 1880 in this sector are
16
17 more consistent, since two companies are reported in each year and the total employment is
18
19 similar in each case at more than 700 workers (U.S. Census Office 1872, 682; U.S. Census
20
21 Office 1883a, 272). Yet the absence of any data on the city's railroad shops in either year
22
23 misrepresents the total employment contribution of the railroad-related manufacturing sector
24
25 and inter-city comparison with Cleveland would be invalid owing to the inconsistency in the
26
27 car manufacturing data there.
28
29

30 31 **Oil-Refining in Cleveland** 32

33
34 Refining of crude oil developed in the city from the early 1860s onwards (Annals of
35
36 Cleveland, 1864, 152), though there had previously been a fledgling industry manufacturing
37
38 oil from coal (Table 6). By the time of the 1870 census, crude refining was an established and
39
40 successful industry. It appears in the printed tabulation, slightly confusingly labelled as
41
42 'coal-oil, rectified', apparently with 16 refineries employing about 200 people and more than
43
44 \$500,000 in capital. The same total number of 16 refineries can be identified in the
45
46 manuscript census, so no major problems have arisen in the classification and tabulation
47
48 process, though the employment totals based on the two sources appear to differ to a limited
49
50 degree (Table 7). Following a now familiar story, the number of companies enumerated does
51
52 not match that contained in city directories or reported by Bradstreet and the Board of Trade,
53
54 indicating approximately a 60% under-enumeration (Cleveland directory 1866, 1871;
55
56
57
58
59
60

1
2
3 Mercantile Agency 1867; Bradstreet's 1870). Comparison of data on capital invested
4 suggests a more serious under-counting for that variable, since it is the smaller companies
5 that are best represented in the manuscript schedules. For example, as early as 1866 the
6 Board of Trade reported nearly 50 refineries 'in more or less successful operation' and an
7 invested capital of \$3 million (Cleveland Board of Trade for 1866, 27), the latter figure
8 nearly 6 times higher than the later census total, even though the industry had grown
9 substantially in the intervening period. By 1870, comparison of sources allows more than 40
10 refineries to be identified by name (Bradstreet's 1870; Cleveland directory 1871). Most
11 importantly, Rockefeller's recently incorporated Standard Oil Company, with a paid up
12 capital of \$1 million, a refining capacity of at least 2500 barrels per day and 100-200
13 employees (cf. Joblin, 1869, 422), is noticeable by its absence from the manuscript schedule,
14 so once again there is incontrovertible evidence of bias towards smaller firms.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 To correct this bias, as far as surviving sources allow, the following procedure was
31 adopted. All the refineries listed in both the 1866/7 and 1871 city directories, but not present
32 in the manuscript census, were identified (Cleveland directory 1866, 1871). These were
33 matched, where possible, to extant newspaper reports dating from 1866, which provided
34 employment figures for a number of firms (Cleveland Daily Leader, January 1866 issues).
35 This is far from perfect, as a basis for 1870 employment estimation, given the time interval
36 involved, but it should provide reasonably conservative baseline figures, since the output of
37 the refining sector was more than three times larger by the census year than it had been in
38 1866 (Cleveland Board of Trade for 1870, 25). Aside from any refineries that were later to be
39 merged into Standard Oil, there were eight firms that matched the relevant criteria, with a
40 total of 92 employees, or an average of 11.5 each. Comparison of names of other firms
41 derived both from the census and the directories, with the extensive but not completely
42 comprehensive Bradstreet's list, revealed a further 18 oil refining companies in 1870, for
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 which no newspaper derived employment figures are extant (Bradstreet's 1870). The average
4
5 employment figure derived above (11.5) was therefore applied to these, yielding a further 207
6
7 workers. The final estimate was obtained by summation of the estimated employment in
8
9 these two refinery groups, the numbers from the manuscript census (193) and a midpoint
10
11 value (150) from the estimated employment range for Standard Oil given in the previous
12
13 paragraph. This produced the final figure of $(92+207+193+150) = 642$, which appears in
14
15 Table 7. This is more than three times the published census total.
16
17
18

19
20 The 1880 enumeration provides multiple further twists to the problem of identifying
21
22 trends in this sector. The relevant category is 'oil, lubricating', which can be found in the
23
24 published tabulation labeled 'selected statistics of manufactures' (U.S. Census Office 1883a,
25
26 322-3). According to this table, the industry has now only 5 players, 125 employees and a
27
28 still modest capitalization of \$656,000. However, close comparison of these figures with the
29
30 manuscript schedules, the published Special Census Report on petroleum and the nearest
31
32 extant Mercantile Agency tabulation from July 1879 (Mercantile Agency 1879; U.S. Census
33
34 Office 1884, 186-191), reveals a number of important issues. Firstly, 'oil, lubricating' in
35
36 1880 is not comparable to the category of 'coal-oil, rectified' in the 1870 census. Marginalia
37
38 from the 1880 schedules distinguish firms counted under the former heading from two other
39
40 categories, one referred to as 'petroleum refiners', the other concerned with varnish
41
42 manufacture. *Neither of these latter two categories is reported in the published 1880*
43
44 *manufacturing census tabulation.* Secondly, the Standard Oil Company was categorized as a
45
46 petroleum refiner; hence it does not appear in the published tabulation. Since it reported a
47
48 capitalization of \$3.5 million, an average workforce of 1935 people (including 5 women) and
49
50 a value of product in excess of \$5.65 million, this absence leaves a major gap in the published
51
52 census record of Cleveland manufacturing activity. Thirdly, the marginalia also indicate that
53
54 some of the data were to be transferred to the Special Agent dealing with the petroleum
55
56
57
58
59
60

1
2
3 industry (U.S. Census Office 1884). His report names a total of 86 refineries across the
4
5 USA, giving the city where they were located, but only aggregate national figures on
6
7 employment and other census variables are provided, so the employment in the sector in
8
9 specific cities cannot be identified. Examination of the refinery names, in the light of the lists
10
11 derived from other sources, also reveals that some of the concerns tabulated under the 'oil,
12
13 lubricating' category in the manufacturing census have now been transferred to the refiners
14
15 category in the Special Report, although the latter makes no mention of this. Using the two
16
17 sources in conjunction, without detailed cross-referencing to the manuscript schedules, would
18
19 therefore result in double-counting, e.g. of Merriam and Morgan, well-known in Cleveland as
20
21 combined paraffin refiners and lubricating oil producers (Cleveland Board of Trade for 1870,
22
23 26), who appear in both tabulations.
24
25
26

27
28 A different 1881 statistical compilation, made at the behest of the Cleveland Board of
29
30 Trade, managed to circumvent the reporting limitations of the published census in relation to
31
32 petroleum refineries, seemingly by obtaining access to advance sheets of the original returns.
33
34 It therefore published aggregate details of the same six refineries as were listed in the
35
36 manuscript census, judging by the almost exact correspondence in the employment figures
37
38 (Cobleigh 1881, 115, 123) (Table 8). Further examination of the Mercantile Agency list from
39
40 1879 suggests this was not the complete picture, since a further nine small refineries are
41
42 found there, making 15 in all (Mercantile Agency 1879). To provide a more complete
43
44 employment estimate, the average number of workers (18.6) employed by the five small
45
46 petroleum refineries recorded in the census, excluding Standard Oil, was multiplied by the
47
48 number of missing refineries to yield the final estimate of employment in this sector of 2195.
49
50 A similar approach was adopted for the lubricating oil sub-sector, to account for a missing
51
52 firm found only in the Bradstreet list. The resulting total of both the lubricating oil and
53
54
55
56
57
58
59
60

1
2
3 petroleum refinery sub-sectors, which avoids double-counting, is 2351. It would be slightly
4
5 larger again; if any of the varnish manufacturers were included.
6
7

8
9 Just as for the car manufacturing sector, this final figure is dramatically greater than
10 that which is obtained by using the population census IND50US code 476 (the OCC and
11 OCC50US variables lack specific codes for refinery workers). Using this industry code, only
12 447 workers can be identified in Cleveland. Closer examination of their individual entries in
13
14 both the coded dataset and the occupational transcriptions reveals that all have an
15
16 occupational description that refers to employment in an oil works of some kind, and 353 of
17
18 them have an OCC code of 74990 for 'specified chemical worker'. Most of the remaining
19
20 workers given a IND50US code of 476 are in various management or clerical positions, but
21
22 most importantly, only 14 of the 447 total have an OCC code of 81500 for 'coopers, hoop
23
24 makers and benders'. It is clear, therefore, that these workers have overwhelmingly *not* been
25
26 attributed to the oil industry in the population census, since there are actually a total of 764
27
28 workers with an OCC code of 81500 in the city.
29
30
31
32
33
34

35
36 The changing relationship of coopers to the refining industry helps explain both the
37
38 mismatch between the 1880 manufacturing and population census derived figures for the
39
40 sector, and the very marked rise in Standard Oil's workforce since the end of the 1860s. It
41
42 also provides evidence of the earlier general observation that changes in company
43
44 organization could markedly affect the distribution of employment between sectors. It was
45
46 virtually a truism that the oil industry had an insatiable demand for barrels, and in 1869 it was
47
48 claimed that, in addition to his refinery workforce, Rockefeller was indirectly giving
49
50 employment to another 7-800 men around the city making barrels (Joblin, 1869, 422). This is
51
52 supported by the published 1870 census, which records 63 independent coopers, with a total
53
54 workforce of just over 600 (U.S. Census 1872, 712). The larger of these firms, in particular,
55
56 indicated specifically on the manuscript schedules that they were making barrels for the oil
57
58
59
60

1
2
3 industry (U.S. Census 1870a). However, during the 1870s Standard Oil removed its
4
5 backward linkages to the cooperage sector and began manufacturing its own barrels, rather
6
7 than buying them from other suppliers. An 1878 report describes both different types of
8
9 barrel-making at the Standard works and wages paid to coopers. The latter are clearly stated
10
11 to be part of the company workforce by this date (Ohio, Bureau of Labor Statistics 2nd
12
13 Annual Report for 1878, 220). Not surprisingly, the 1880 tabulation then suggests that the
14
15 cooperage sector has shrunk dramatically, since only 13 establishments are recorded,
16
17 employing less than 200 people. Thus, many hundreds of workers had disappeared from the
18
19 sector between 1870 and 1880 according to the published manufacturing census, but the
20
21 evidence from the population census shows otherwise (764 coopers). Had Standard Oil been
22
23 tabulated in accordance with the data in the manuscript schedule, many of these same missing
24
25 coopers would, of course, have bolstered the oil-refining sector instead.
26
27
28
29

30
31 Accounting for the missing coopers is not sufficient to explain the tenfold increase in
32
33 employment at Standard Oil during the 1870s. The 1869 report also makes reference to the
34
35 expenditure of \$20-25,000 a year by Rockefeller to pay tradesmen not on the company
36
37 payroll, for repairs around the refinery complex (Joblin 1869, 422). It is very likely,
38
39 therefore, that subsequently, much of this work was moved 'in-house', in the same way as for
40
41 the coopers, to produce the same type of generic worker attribution problem, as found on the
42
43 railroads, i.e. by 1880 there were many tradesmen working for Standard Oil, who did not
44
45 declare their workplace affiliation.
46
47
48

49
50 Although such detailed examples quickly become tiresome, the numerical magnitudes
51
52 of the errors, omissions and inconsistencies they represent constitute serious problems for
53
54 accurate interpretation of the data. These problems are effectively insurmountable, using
55
56 only the census itself, as comparison with a variety of non-census sources is required to
57
58 identify them and estimate their magnitudes, on a sector by sector and location by location
59
60

1
2
3 basis. This effectively renders the application of general correction or weighting factors
4
5 infeasible, unless and until future work finds more regularities in other sectors than has been
6
7 found in those investigated here.
8
9

10
11 A number of distinguished commentators have regrettably been caught by these
12
13 problems, without realizing it. For example, in his wider study of midwestern
14
15 industrialization, Meyer makes, among others, inter-censal comparisons between 1870 and
16
17 1880 of employment concentration in his sector labeled 'petroleum and coal' (Meyer 1989,
18
19 930), but these are not tenable in the Cleveland context. The same kind of underreporting of
20
21 oil-refining is apparent elsewhere in the census, for example in the Pittsburgh statistics for
22
23 1880, so this finding is of more general applicability (U.S. Census Office, 1883a, 338). It
24
25 also affects the comparisons made by Lamoreaux and Levenstein (2008, Table 1, 35) and
26
27 Kim (1999, 5-7) falls into the same data trap in his analysis of the role of factor endowments
28
29 in industrial development. The latter uses the 1880 'petroleum and coal' manufacturing
30
31 sectors from the census as part of his panel regression analysis by state, although strangely,
32
33 while he employs the census data on coal mining, he does not appear to use the
34
35 corresponding data from the census special report on petroleum. This provides equivalent
36
37 figures for oil wells, which might then have been compared with refinery output, according to
38
39 the logic of his analysis (U.S. Bureau of the Census 1884, 1886). Thus his manufacturing
40
41 data are incorrect for the petroleum sector (and indeed for transportation) judging by his
42
43 Table 2. Yet even if they had been present and correct, his logic would have fallen foul of
44
45 the fact that intra-state comparisons are meaningless for oil-refining at that date, since the
46
47 vast majority of the crude oil originated in Pennsylvania, but was refined either in Cleveland
48
49 or New York.
50
51
52
53
54
55
56
57
58
59
60

Conclusions

A detailed investigation of the quality of the US Manufacturing Census 1860-1880 has been undertaken using case study examples from several sectors in the fast-growing city of Cleveland, OH, and comparisons with other industrial cities in surrounding states. These have confirmed that the enumeration limitations first reported by the 1870 census are not worse case scenarios, but rather the norm, at least for these examples. While previous writers have stressed clerical errors and bureaucratic tampering in the tabulation of data for the printed reports, these have not been found to be the most serious or pervasive issues, in terms of potential analytical use of the data. Instead, several other issues have been found to be the most worrying. The first is varying, but substantial levels of under-enumeration, which can easily exceed 50%, coupled with a clear tendency towards the omission of the larger companies across multiple sectors, (and there is further evidence from the primary iron manufacture sector not reported here). This omission affects the manuscript schedules (and hence any digital samples derived from them) almost as much as the printed reports. The second major concern is that inconsistent definitions *between* census years, and inconsistent application of prevailing definitions *within* census years of what constituted manufacturing has led to the non-enumeration of those major portions of the railroad industry, which were involved with the construction and maintenance of locomotives and rolling stock. Given the importance of forward and backward linkages associated with the railroad industry during the period (Fishlow 1965), this has major implications for our understanding of the interrelationships between transportation and manufacturing development. Similar inconsistencies, compounded by changes in industrial structure, have been found to affect the oil refining industry. Correcting for these deficiencies shows that the railroads were in fact the leading employment sector in Cleveland in 1860 and 1870, only slipping into second place behind primary metal manufacture in 1880, while refining and associated cooperage

1
2
3 was the third most important sector in both 1870 and 1880 (Healey 2014). Combined
4
5 railroad and refining-related employment constituted an estimated 25.6% of total railroad and
6
7 manufacturing workers in the city in 1870 and 22.1% of the much larger equivalent
8
9 employment base in 1880 (cf. Table 1 in Healey 2014). Using the manufacturing census
10
11 alone, would have produced corresponding, but erroneous figures of 11.3% and 0.4%.
12
13
14 Thirdly, both the original compilers of the census publications and a number of researchers
15
16 who have used their tabulations subsequently, have paid too little attention to ensuring, as far
17
18 as possible, that classificatory categories adopted in different census years were sufficiently
19
20 similar to permit of a reasonable degree of inter-censal comparability. In consequence, the
21
22 only element of consistency seems to be in terms of the bias against large companies, but
23
24 even here, the presence or absence of individual concerns in succeeding censuses may result
25
26 in tabulated employment and capital levels, which wrongly appear to fluctuate by many
27
28 hundreds/thousands or millions respectively.
29
30
31

32
33 Such large deviations from accurate enumeration of manufacturing activity are likely
34
35 to be of the same or greater order than the kinds of statistical effects found in quantitative
36
37 studies using these data, which is hardly an incentive to pursue this kind of investigation.
38
39 Equally, large enumeration errors can very easily lead to quite false inferences from limited
40
41 use of census totals in more qualitative studies. In an era where research resources are
42
43 stretched, such findings will be unwelcome to all and wholly unpalatable to many.
44
45
46 Nevertheless, the only reasonable conclusion to draw is that the manufacturing census can
47
48 *only* safely be used as one of a number of types of partial evidence of economic activity, each
49
50 of which has to be triangulated against the others, in a manner perhaps similar to that
51
52 undertaken earlier in this paper. The manuscript census has value in providing details of
53
54 named concerns at specific dates, but neither it nor the published reports provide reliable,
55
56 consistent and comparable data on the relative size of manufacturing sectors within or
57
58
59
60

1
2
3 between censuses or on their aggregate inputs and outputs. Development of new
4
5 manufacturing datasets that could be used with confidence in, for example, large statistical
6
7 panel regressions, would, in the present writer's view, be an undertaking of enormous
8
9 proportions. Much more work than that reported here just for Cleveland would have to be
10
11 repeated several hundred-fold across the major urban centers/counties of the United States,
12
13 before acceptable estimates of manufacturing activity could be derived, if indeed the
14
15 necessary evidence to make such estimates could be found in all cases. As a first step in the
16
17 right direction, work has now commenced to identify the locations of all railroad shops and
18
19 their approximate establishment dates across the American Manufacturing Belt. This will
20
21 provide a first yardstick for systematic, wide area evaluation of the full extent of one of the
22
23 more glaring manufacturing census deficiencies identified in this study.
24
25
26
27
28
29
30

31 **References**

32
33
34 Alexander, J.G. 1982. City directories as "ideal censuses": Slovak immigrants and
35
36 Pittsburgh's early twentieth-century directories as a test case. *The Western Pennsylvania*
37
38 *Historical Magazine* 65(3): 203-220.
39

40
41
42
43
44 Annals of Cleveland. 1850-1876, transcription 1936 -. *Annals of Cleveland, Works Progress*
45
46 *Administration in Ohio. Years 1850-1876, 27 vols, (approx. 17,000 pages of typescript).*
47
48

49 Cleveland: WPA, District Four.
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Annual Report of the Secretary of Internal Affairs. 1877. *Annual Report of the Secretary of*
4
5 *Internal Affairs of the Commonwealth of Pennsylvania, Part III, Industrial Statistics 1876-77.*
6
7 Harrisburg: Lane S. Hart, State Printer
8
9

10
11
12
13 Annual Report of the Secretary of Internal Affairs. 1881. *Annual Report of the Secretary of*
14
15 *Internal Affairs of the Commonwealth of Pennsylvania, Part III, Industrial Statistics 1879-80.*
16
17 Harrisburg: Lane S. Hart, State Printer
18
19

20
21
22
23 Atack, J. 1987. Economies of scale and efficiency gains in the rise of the factory in America,
24
25 1820-1900. In *Quantity and quiddity: Essays in U.S. economic history*, edited by P. Kilby,
26
27 286-335. Middletown, CN: Wesleyan University Press.
28
29

30
31
32
33 Atack, J. and F. Bateman. 1992. How long was the workday in 1880? *Journal of Economic*
34
35 *History* 52:129-160.
36
37

38
39
40
41 Atack, J. and F. Bateman. 1999. Nineteenth century U.S. industrial development through the
42
43 eyes of the census of manufactures: a new resource for historical research. *Historical*
44
45 *Methods*, 32:177-188.
46
47

48
49
50
51 Atack, J., F. Bateman and R. A. Margo. 2002. Part-year operation in nineteenth-century
52
53 American manufacturing: evidence from the 1870 and 1880 censuses. *Journal of Economic*
54
55 *History* 62:792-809.
56
57

1
2
3
4
5
6 Atack, J., M. Haines and R. A. Margo. 2011. Railroads and the rise of the factory: evidence
7
8 for the United States 1850-1870. In *Economic evolution and revolution in historical time*,
9
10 edited by P. W. Rhode, D. F. Weiman and J. L. Rosenbloom, 162-179. Stanford: Stanford
11
12 University Press.
13

14
15
16
17
18
19 Bateman, F., T. Weiss, and J. Atack. n.d. *Samples from the manuscript censuses of*
20
21 *manufactures 1850-1880*. Available on-line at
22
23 <http://www.vanderbilt.edu/Econ/faculty/Atack/atackj.htm>. Accessed July 4, 2012 and at
24
25 <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4048> (1850-70 samples) and at
26
27 www.icpsr.umich.edu/icpsrweb/ICPSR/studies/9384 (1880 census sample). Accessed Feb 14,
28
29 2015.
30
31

32
33
34
35
36 Boston and Maine Railroad. 1861. *Report of the directors of the Boston and Maine Railroad*
37
38 *to the stockholders. Wednesday, Sept. 11, 1861*. Boston: Henry W. Dutton and Son.
39
40

41
42
43
44 Bradstreet's. 1870. *Bradstreet's commercial reports embracing the bankers merchants*
45
46 *manufacturers and others in the United States and the Dominion of Canada, Vol. XXVII, July*
47
48 *6, 1870*. New York: J.M. Bradstreet and Son.
49
50

51
52
53
54
55 Chandler, A. D. 1962. *Strategy and structure: chapters in the history of the industrial*
56
57 *enterprise*. Cambridge, MA: M.I.T. Press.
58
59
60

1
2
3
4
5
6 Chandler, A.D. 1977. *The visible hand: the managerial revolution in American business.*

7
8 Cambridge, MA: Harvard University Press.
9

10
11
12
13
14 Cleveland and Toledo Railroad. 1854 -. *Annual report for the year 1853/4 etc.*
15
16

17
18
19
20 Cleveland Board of Trade. 1866 -. *Annual report of the trade, commerce and manufactures of*
21
22 *the city of Cleveland, for the year 1865 etc. Reported to the Board of Trade by J. D.*
23

24 *Pickands.* Cleveland: Fairbanks, Benedict and Company.
25
26

27
28
29
30
31 Cleveland, Columbus and Cincinnati Railroad. 1852 -. *Annual report for the year 1851 etc.*
32
33

34
35
36
37 Cleveland Daily Leader (October 1865 – August 1866) Newspaper online version available at
38
39 <http://chroniclingamerica.loc.gov/lccn/sn85042437/issues/1866/> (Accessed January 2, 2015).
40
41

42
43
44
45 Cleveland Directory. 1861. *J. H. Williston and Co.'s directory of the city of Cleveland to*
46
47 *which is added a business directory for 1861-62.* Cleveland: Ben Franklin Print.
48
49

50
51
52
53 Cleveland Directory. 1866. *Cleveland Leader city directory, for 1866-67, together with a*
54
55 *business directory.* Cleveland: Cleveland Leader Publishers.
56
57

1
2
3
4
5
6 Cleveland Directory. 1871. *W. S. Robison and Co.'s Cleveland directory 1871-2*, A. Bailey,
7
8 *Compiler*. Cleveland: W.S. Robison and Co.
9

10
11
12
13
14 Cleveland Directory. 1880. *The Cleveland directory for the year ending June 1881, Vol X*.
15
16 Cleveland: Cleveland Directory Co.
17

18
19
20
21
22 Cleveland Morning Leader (June 1858-April 1865) Newspaper online version available at
23 <http://chroniclingamerica.loc.gov/lccn/sn83035143/issues/1858/> (Accessed December 28
24
25
26
27 2014).
28

29
30
31
32
33 Cobleigh, N.S. 1881. *The manufactures, trade and commerce of Cleveland, 1880-81*.
34
35 Cleveland: Short and Forman Printers.
36

37
38
39
40
41 Cootner, P.H. 1963. The role of the railroads in United States economic growth. *Journal of*
42
43 *Economic History*, 23(4): 477-521.
44

45
46
47
48
49 Cram, G. F, J. H. Beers and J. Q. A. Bennett. 1892. *Atlas of Cuyahoga County and the city of*
50
51 *Cleveland Ohio*. Chicago: G.F. Cram and Company.
52

53
54
55
56
57
58 De Geer, S. 1927, The American Manufacturing Belt. *Geografiska Annaler*, 9: 233-359.
59

1
2
3
4
5
6 Delle Donne, C. R. 1973. Federal census schedules, 1850-1880: primary sources for
7
8 historical research. *National Archives and Records Service Reference Information Paper No.*
9
10 67. Washington: General Services Administration.
11

12
13
14
15
16 DLWRR Coal Department. 1860, 1870. *Delaware, Lackawanna and Western Railroad Coal*
17
18 *Department Records, Payrolls June 1860, 1870.* Lackawanna Historical Society, Scranton,
19
20 PA.
21

22
23
24
25
26
27 Encyclopedia of Cleveland History. 1997. *Collinwood railroad yards.* On-line at
28
29 <http://ech.case.edu/ech-cgi/article.pl?id=CRY>, accessed 15 April, 2012.
30
31

32
33
34
35 Fels, R. 1959. *American business cycles, 1865-1897.* Chapel Hill: University of North
36
37 Carolina Press.
38

39
40
41
42
43 Fishlow, A. 1965. *American railroads and the transformation of the ante-bellum economy.*
44
45 Cambridge, MA: Harvard University Press.
46
47

48
49
50
51 Fogel, R. W. 1964. *Railroads and American Economic Growth: Essays in Econometric*
52
53 *History.* Baltimore: Johns Hopkins Press.
54
55

1
2
3 Goldstein, S. 1954. City directories as sources of migration data. *American Journal of*
4
5 *Sociology* 60(2): 169-176.
6
7

8
9
10
11 Gregory, I and Ell, P. S. 2007. *Historical GIS: technologies, methodologies, and scholarship.*
12
13
14 Cambridge: Cambridge University Press.
15
16

17
18
19
20 Healey R. 2011. A full-scale implementation of the NAPP 1880 U.S. Census data set using
21
22 dimensional modeling and data-warehousing technology *Historical Methods: A Journal of*
23
24 *Quantitative and Interdisciplinary History*, 44 (2), 95-105.
25
26

27
28
29
30 Healey R., Thomas W. and Lahman K. 2013. Railroads and regional labor markets in the
31
32 mid-nineteenth-century United States: a case study of the Baltimore and Ohio Railroad
33
34 *Journal of Historical Geography*, 41, 13-32.
35
36

37
38
39
40 Healey, R.G. 2014. Railroads, factor channelling and increasing returns: Cleveland and the
41
42 emergence of the American manufacturing belt. *Journal of Economic Geography*, Advance
43
44 *Online Publication:*
45
46 <http://joeg.oxfordjournals.org/content/early/2014/07/14/jeg.lbu011.abstract>.
47
48
49

50
51
52
53 Healey R. G. 2015. A New Occupational/Industrial Coding System for 19th Century U.S.
54
55 Heavy Industrial Workers. Manuscript submitted to *SAGE Open*.
56
57
58
59
60

1
2
3
4
5
6 Hirsch, S. E. 1978. *Roots of the American working class: the industrialization of crafts in*
7
8 *Newark, 1800-1860*. Philadelphia: University of Pennsylvania Press.
9

10
11
12
13
14 Hopkins, G. M. 1858. *Map of Cuyahoga County, Ohio*. Philadelphia: S.H. Matthews.
15
16

17
18
19
20 Inwood, K. 1995. The census representation of Canadian industry, 1871-91. *Histoire*
21
22 *sociale/Social History*, 28, 347-74.
23
24

25
26
27
28
29 Jenks, L. H. 1944. Railroads as an economic force in American development. *Journal of*
30
31 *Economic History*, 4:1-20.
32
33

34
35
36
37 Joblin, M. 1869. *Cleveland, past and present: its representative men: Comprising*
38
39 *biographical sketches of pioneer settlers and prominent citizens, with a history of the city.*
40
41
42 Cleveland: Maurice Joblin.
43
44

45
46
47
48 Kim, S. 1995. Expansion of markets and the geographic distribution of economic activities:
49
50 trends in U.S. regional manufacturing structure, 1860-1987. *Quarterly Journal of Economics*,
51
52 110:881-908.
53
54

1
2
3 Kim, S. 1999. Regions, resources, and economic geography: sources of U.S. regional
4 comparative advantage, 1880-1987. *Regional Science and Urban Economics*, 29: 1-32.
5
6

7
8
9
10 Klein, A. and Crafts, N. 2011. Making sense of the Manufacturing Belt: determinants of U.S.
11 industrial location, 1880-1920. *Journal of Economic Geography, Business and Economic*
12 *Geography Virtual Issue*: 1-33, doi: 10.1093/jeg/lbr023.
13
14

15
16
17
18
19
20 Knights, P.R. 1969. City directories as aids to ante-bellum urban studies: a research note.
21
22 *Historical Methods Newsletter* 2(4): 1-10.
23
24

25
26
27
28 Lake Shore and Michigan Southern Railway. 1871-. *Annual report for 1870 etc.*
29
30

31
32
33
34 Licht, W. 1983. *Working for the railroad: The organization of work in the nineteenth*
35 *century*. Princeton: Princeton University Press.
36
37

38
39
40
41
42 Long, J.H. and Tuck Sinko, P. (Eds.). 2010. *Atlas of Historical County Boundaries*.
43
44 <http://publications.newberry.org/ahcbp/project.html>. (Accessed Dec. 28, 2014).
45
46

47
48
49
50
51 Mercantile Agency. 1867. *The Mercantile Agency reference book, (and key), containing*
52 *ratings of the merchants, manufacturers and traders generally, throughout the United States.*
53
54
55
56 *Vol. 22, January 1867*. New York: R.G. Dun and Co.
57
58
59
60

1
2
3
4
5
6 Mercantile Agency. 1879. *The Mercantile Agency reference book, (and key), containing*
7
8 *ratings of the merchants, manufacturers and traders generally, throughout the principal*
9
10 *cities of the United States and Canada. Vol. 45, July 1879.* Philadelphia, Baltimore etc.: R.G.
11
12 Dun and Co.

13
14
15
16
17
18 Meyer, D. R. 1983. Emergence of the American Manufacturing Belt: an interpretation.
19
20 *Journal of Historical Geography*, 9: 145-174.
21
22

23
24
25
26 Meyer, D. R. 1989. Midwestern industrialization and the American Manufacturing Belt in
27
28 the nineteenth century. *Journal of Economic History*, 49:921-37.
29
30

31
32
33
34 Meyer, D. R. 1990. The rise of the industrial metropolis: the myth and the reality. *Social*
35
36 *Forces*, 68:731-752.
37
38

39
40
41
42 Michigan Central Railroad. 1870. *Annual report for 1869/70.* Boston: Wright and Potter.
43
44
45

46
47
48
49 Minnesota Population Center. 2004. *National Historical Geographic Information System*
50
51 *(NHGIS): Pre-release Version 0.1.* Minneapolis, Minnesota: University of Minnesota.
52
53 <http://www.nhgis.org>.
54
55
56
57
58
59
60

1
2
3 Minnesota Population Center. 2008. *North Atlantic Population Project: Complete Count*
4
5 *Microdata. Version 2.0 [Machine-readable database]*. Minneapolis: Minnesota Population
6
7 Center. (<http://www.nappdata.org>).
8
9

10
11
12
13 Niemi, A. W. 1974. *State and Regional Patterns in American Manufacturing, 1860-1900*.
14
15 London : Greenwood Press.
16
17

18
19
20
21 North, S. N. D. 1890. The annual statistics of manufactures (Massachusetts) 1886 : the annual
22
23 statistics of manufactures 1888. *Publications of the American Statistical Association* 2:22-29.
24
25
26

27
28
29
30 North, S. N. D. 1899. Manufactures in the federal census. *Publications of the American*
31
32 *Economic Association, New Series* 2:257-302.
33
34
35

36
37
38 Ohio, Bureau of Labor Statistics. 1877 -. *First etc. annual report of the Bureau of Labor*
39
40 *Statistics made to the General Assembly of Ohio*. Columbus: State Printer.
41
42
43

44
45
46 Ohio, Commissioner of Railroads. 1868 -. *Annual report of the Commissioner of Railroads*
47
48 *and Telegraphs to the Governor of the State of Ohio for the year 1867 etc*. Columbus: State
49
50 Printer.
51
52
53
54
55
56
57
58
59
60

1
2
3 Ohio, Commissioner of Statistics. 1857 -. *First etc. annual report of the Commissioner of*
4
5 *Statistics to the General Assembly of Ohio.* Columbus: State Printer.
6
7
8
9

10
11 Perrin, W. H., J. H. Battle and W. A. Goodspeed. 1881. *History of Crawford County and*
12
13 *Ohio.* Chicago: Baskin and Battey.
14
15
16
17
18

19
20 Price, W.W. 1933. *We Have Recovered Before! A Comparison of the Present*
21
22 *Depression with the Major Depressions of the Past Century: 1837, 1857, 1873, and*
23
24 *1893.* New York: Harper and Brothers.
25
26
27
28
29

30 Robison, W. S. 1887. *History of the city of Cleveland: its settlement, rise and progress.*
31
32 Cleveland: Robison and Cockett.
33
34
35
36
37

38 Rose, W.G. 1950. *Cleveland: The making of a city.* Cleveland: World Publishing Company.
39
40
41
42
43

44 Stuart, W. M. 1898. Official statistics. *American Journal of Sociology* 3:622-630.
45
46
47
48
49

50 Swineford, A. P. 1876. *History and review of the copper, iron, silver, slate and other*
51
52 *material interests of the south shore of Lake Superior.* Marquette: the Mining Journal.
53
54
55
56
57
58
59
60

1
2
3 U.S. Census Office. 1854. *Statistical view of the United States*. Washington: Beverley
4
5 Tucker.
6
7
8
9

10
11 U.S. Census Office. 1865. *Manufactures of the United States in 1860; Compiled from the*
12
13 *original returns of the Eighth Census under the direction of the Secretary of the Interior.*
14
15 Washington: GPO.
16
17
18
19

20
21 U.S. Census Office. 1870a. *Manuscript returns of the 1870 Census of Manufactures*. State
22
23 of Ohio (part). Microcopy T1159 Roll 45. Washington: NARA.
24
25
26
27
28
29

30 U.S. Census Office. 1870b. *Ninth census, United States, 1870. Instructions to assistant*
31
32 *marshals. Act of May 23, 1850*. Washington: GPO.
33
34
35
36
37

38 U.S. Census Office. 1872. *Vol. III, The statistics of the wealth and industry of the United*
39
40 *States compiled from the original returns of the Ninth Census (June 1, 1870) under the*
41
42 *direction of the Secretary of the Interior by Francis A. Walker*. Washington: GPO. Digital
43
44 version available from <https://www.census.gov/prod/www/decennial.html>. Viewed March 3,
45
46 2015.
47
48
49
50
51
52

53 U.S. Census Office. 1880. *Manuscript returns of the 1880 Census of Manufactures*. State of
54
55 Ohio (part). Microcopy T1159 Roll 93. Washington: NARA.
56
57
58
59
60

1
2
3
4
5
6 U.S. Census Office. 1883a. *Vol. 2, Report on the manufactures of the United States at the*
7
8 *Tenth Census (June 1, 1880)*. Washington: GPO.
9

10
11
12
13
14 U.S. Census Office. 1883b. *Vol. 4, Report on the agencies of transportation in the United*
15
16 *States including the statistics of railroads, steam navigation, canals, telegraphs, and*
17
18 *telephones*. Washington: GPO.
19

20
21
22
23
24 U.S. Census Office. 1884. *Vol. 10, Production, technology and uses of petroleum and its*
25
26 *products by S.F. Peckham; The manufacture of coke by J.D. Weeks; Building stones of the*
27
28 *United States, and statistics of the quarry industry for 1880*. Washington: GPO.
29
30
31

32
33
34
35 U.S. Census Office. 1886. *Vol. 15, Report on the mining industries of the United States*
36
37 *(exclusive of the precious metals) with special investigation into the iron resources of the*
38
39 *Republic and into cretaceous coals of the Northwest, by R. Pumpelly*. Washington: GPO.
40
41
42

43
44
45 U.S. Census Office. 1888. *Vol. 22, Report on power and machinery employed in*
46
47 *manufactures, embracing statistics of steam and water power used in the manufacture of iron*
48
49 *and steel, machine tools and wood-working machinery, wool and silk machinery etc.*
50
51
52 Washington: GPO.
53
54
55
56
57
58
59
60

1
2
3 U.S. Census Office. 1895a. *Vol. VI, Report on manufacturing industries in the United States*
4
5 *at the Eleventh Census: 1890*. Washington: GPO.
6
7

8
9
10
11 U.S. Census Office. 1895b. *Vol. XIV, Report on transportation business in the United States*
12
13 *at the Eleventh Census: 1890, Part I, transportation by land*. Washington: GPO.
14
15

16
17
18
19 U.S. Census Office. 1900. *Twelfth census of the United States. June 1, 1900. Instructions to*
20
21 *enumerators. Under the provisions of the acts of Congress approved March 3, 1899, and*
22
23 *February 1, 1900*. Washington: GPO.
24
25

26
27
28
29 U.S. Census Office. 1902. *Vol. VIII, Manufactures: Part II, states and territories*.
30
31 Washington: GPO.
32
33

34
35
36
37
38 Walkowitz, D.J. 1978. *Worker city, company town: Iron and cotton-worker protest in Troy*
39
40 *and Cohoes, New York, 1855-84*, Urbana: University of Illinois Press.
41
42

43
44
45
46 Walsh, M. 1970. The census as an accurate source of information: the value of mid-
47
48 nineteenth century manufacturing returns. *Historical Methods Newsletter*, 3(4):3-13.
49
50

51
52
53
54
55 Warren, K. 1973. *The American steel industry 1850-1950: A geographical interpretation*.
56
57 Oxford: Clarendon Press.
58
59

1
2
3
4
5
6 Wilson, M.R. 2005. *Railroad Supply Industry*, entry in on-line Encyclopedia of Chicago
7
8 History <http://www.encyclopedia.chicagohistory.org/pages/2387.html>. Accessed May 5,
9
10 2012.
11

12
13
14
15
16 Wright, C. D. 1900. *The history and growth of the United States census prepared for the*
17
18 *Senate committee on the census*. Washington: GPO.
19
20
21
22
23
24
25
26

27
28 **List of Tables**
29

30
31

32 Table 1	Categorisation of Within Census Problems
33 Table 2	Categorisation of Between Census Problems
34 Table 3	Cleveland 1861 City Directory: Employment Breakdown by 35 Occupation
36 Table 4	Cleveland 1861 City Directory: Employment Breakdown by Railroad
37 Table 5	Cleveland 1861 City Directory: Employment Breakdown Summary
38 Table 6	Cleveland 1860/1: Summary of Employment Estimates by Sector
39 Table 7	Cleveland 1870: Summary of Employment Estimates by Sector
40 Table 8	Cleveland 1879/80: Summary of Employment Estimates by Sector

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Enumeration Process		Tabulation Process	
In 1870 Walker Report	Problem Type	Problem Type	References
Y	Manufacturing Census data collection lower priority than Population Census	Clerical errors in processing the manuscript schedules (e.g. FL, OR, CA)	Atack and Bateman (1999); Walsh (1970)
Y	Single visit opportunity to collect all required data	'Massaging' of data by Washington bureaucrats	Atack and Bateman (1999)
Y	Low level of payment per firm enumerated	Printing Errors (e.g. CA)	Atack and Bateman (1999)
Y	Low remuneration encouraged negligent performance of duties	Absence of surviving documentation of general classification methodologies	This paper
Y	Failure to find small firms hidden in basements and interior courtyards	Specific information loss from classifying multi-product firms into single product categories	This paper
N	Lack of co-operation in reporting by large firms leading to small size bias	Omission of large firms reported on manuscript schedules from tabulations	This paper
Y	Branch and main offices or plant sites in different locations leading to non-enumeration (each site believing the other was reporting)	Incomplete national coverage of surviving manuscript schedules hinders checking of aggregate tabulations	This paper; Atack and Bateman (1999)
N	Branch and main offices/plants both reporting leading to double-counting	Incorrect omission of specific standard categories from individual county tabulations	This paper
Y	Understatement of output value by firms close to the \$500 reporting threshold, to avoid enumeration		
Y	Inaccurate and inconsistent reporting by firms of capital invested		
Y	Inaccurate reporting of labor inputs		
N	Geographical (county by county or urban/rural) variations in the severity of the above problems		

Table 1 **Categorisation of Within Census Problems**

Enumeration Process		Tabulation Process	
Problem Type	References	Problem Type	References
Changing universe of eligible firms (changing average firm size)	Hirsch (1978)	Differences in the categories used to subdivide manufacturing	Stuart (1898); This paper
Changing universe of firm composition (effects of vertical integration etc.)	Chandler (1962, 1977); Warren (1973)	Differences in the practice of assigning firms to supposedly the same categories	This paper
Changing definition of manufacturing: inconsistencies in the inclusion/exclusion of specific types of activity	Stuart (1898); This paper	Different levels of small size bias	This paper
Censuses taken at different stages of the business cycle	North (1890)	Geographical inconsistencies in compliance with prevailing classification procedures (i.e. a county may comply in one census but not in another)	This paper
Lack of systematic reporting of 'miscellaneous manufacturing expenses' pre 1890	North(1899)	Changes in county boundaries between censuses	This paper; Long and Tuck Sinko (2010); Gregory and Ell (2007).
Lack of comparability in measurement of machine capacity	North(1899)		
Lack of consistency in measurement of daily and seasonal labor inputs	Atack and Bateman (1992); Atack, Bateman and Margo (2002)		

Table 2 **Categorisation of Between Census Problems**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

	No Railroad Specified Unknown Workplace	Non-Railroad Known Workplace	Vessel Worker	Non-Railroad Identified Total	Railroad No Location Given	Railroad Depot Location	Railroad Other Non -Shop Location	Railroad Shop Location	Railroad Identified Total
Potential Train Men									
Baggage Worker	10	0	1	11	26	1	0	0	27
Brakeman	38	0	0	38	62	0	0	0	62
Conductor	10	1	0	11	46	0	0	0	46
Engineer	74	14	7	95	48	0	0	4	52
Fireman	8	0	0	8	39	0	0	2	41
Category Sub-total	140	15	8	163	221	1	0	6	228
Generic Workers									
Agent	32	44	2	78	27	1	0	0	28
Clerk	326	186	3	515	41	16	4	0	61
Laborer	1218	28	0	1246	6	32	0	2	40
Category Sub-total	1576	258	5	1839	74	49	4	2	129
Generic Tradesmen									
Blacksmith	217	23	5	245	1	9	0	0	10
Boilermaker	39	7	0	46	0	0	0	3	3
Carpenter	381	2	116	499	0	0	1	0	1
Carpenter and Joiner	39	0	0	39	0	0	0	0	0
Joiner	100	0	1	101	1	0	0	1	2
Machinist	158	57	0	215	6	3	0	14	23
Mason	157	0	0	157	0	1	0	0	1
Mechanic	19	1	0	20	1	0	0	2	3
Painter	172	12	1	185	0	0	0	2	2
Tinner	76	0	0	76	0	0	0	0	0

Category Sub-total	1358	102	123	1583	9	13	1	22	45
Overall Sub-total	3074	375	136	3585	304	63	5	30	402
Other Railroad Workers					100	65	0	13	178
Grand Total					404	128	5	43	580

Table 3 Cleveland 1861 City Directory: Employment Breakdown by Occupation

Railroad Company	No. of Identified Employees
Cleveland, Columbus and Cincinnati RR (CCCRR)	182
Cleveland, Painesville and Ashtabula RR (CPARR)	113
Cleveland and Toledo RR (CTRR)	81
Cleveland and Pittsburgh RR	73
Cleveland and Mahoning RR	23
CCCRR/CPARR Jointly	17
CCCRR/CTRR Jointly	3
Other Railroads	16
Railroad not named*	72
Total	580

* Includes one ambiguously abbreviated name

Table 4 Cleveland 1861 City Directory: Employment Breakdown by Railroad

Category	Estimated Unidentified Railroad Men	Identified Railroad Men	Row Total
Train Men	93	228	321
Generic Workers	236	129	365
Generic Tradesmen	440	45	485
Other Railroad Men		178	178
Column Total	769	580	1349
Grand Total			1349

Table 5 Cleveland 1861 City Directory: Employment Breakdown Summary

For Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

	Manuscript Manufacturing Census	Printed Manufacturing Census	Estimated from Newspaper Sources	1861 City Directory Identified Workers Only	Estimated Total Workers based on 1861 City Directory
Railroad Shops	N/A	Not Reported	450-475	43	N/A
Total Railroad Employees	N/A	Not Reported	900	580	1349
Car Manufacture	N/A	30	N/A	14*	N/A
Coal Oil Refining	N/A	3	N/A	8	N/A

* occupation of 'car builder' only, none give any employment location
 N/A data not available

Printed census figures from U.S. Census Office (1865, 447)

For sources of data and calculations pertaining to figures in columns 3-5, see text.

Table 6 Cleveland 1860/1: Summary of Employment Estimates by Sector

	Manuscript Manufacturing Census	Printed Manufacturing Census	Estimated using census and/or other sources
Railroad Shops	960	960	1360
Total Railroad Employees	N/A	Not Reported	2750
Car Manufacture	N/A	10	400*
Oil Refining	193	209	642

* McNairy and Claflen car building and bridge building employment combined total

Printed census figures from U.S. Census Office (1872, 712)

For sources of data and calculations pertaining to final column figures, see text.

Table 7 Cleveland 1870: Summary of Employment Estimates by Sector

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

	Manuscript Manufacturing Census	Printed Manufacturing Census	Estimated using population census (excl. E. Cleveland)	Combined Estimate from Bradstreet and Manufacturing Census	Cleveland Board of Trade Figures
Railroad Shops	Not Reported	Not Reported	1958	N/A	N/A
Total Railroad Employees	N/A	N/A	3832	N/A	N/A
Car Manufacture	213	216	40*	N/A	N/A
Oil, Lubricating	125	125	N/C	156	N/A
Petroleum Refining	2028	Not Reported	447	2195	2027

N/A not available

N/C not calculated - subset of workers cannot be distinguished from other oil refinery workers

* Not including unspecified car shop workers

Printed census figures from U.S. Census Office (1883a, 322-3); see text for sources and calculations for figures in columns 3-5.

Table 8 **Cleveland 1879/80: Summary of Employment Estimates by Sector**