

Online Appendix 3: Data

A. Weeks Worked and Part-Year Work

In all cases except Canada in 1941 and 1951, the weeks worked variables refer to annual weeks worked in the year prior to the census year, and are used to compute the part-year work variables. In 1941-51 Canadian censuses, weeks worked referred to the one year period prior to the June reference week (i.e., June 1940-June 1941 instead of January 1940 to December 1940). Prior to 1980, the U.S. census coverage was individuals 14+, which changed to 16+ in 1980; in the Canadian census, coverage changed from 14+ to 15+ in 1971. In addition, prior to 1971 only wage-earners were asked their weeks worked in Canada (i.e., self-employed and unpaid family workers were excluded). In the cell-level analysis this variable is the proportion of individuals in the gender/industry/region/year cell that work less than 40 weeks.

Weeks worked are available as continuous variables only in the 1980/81 and 1990/91 U.S. and Canadian censuses. In previous years, in each country, weeks worked were categorized as follows: U.S. (all years) 1-13, 14-26, 27-39, 40-47, 48-49, 50-52; Canada (1961, 1971) 1-13, 14-26, 27-39, 40-48, 49-52; Canada (1941, 1951) 1-19, 20-29, 30-39, 40-49, 50-52. In the 1940 U.S. census and the 1941 and 1951 Canadian censuses, individuals who worked part-time had their weeks worked converted into full-time equivalent weeks worked (i.e., an individual working 40 weeks at 20 hours per week would appear in the data as 20 weeks worked). We convert reported weeks worked into an estimated actual weeks worked for these three sets of cells using the hours worked information in the U.S. census. We rely on the assumption that the propensity to work part-time was the same in Canada as it is was in the U.S. Individuals in the 1940 U.S. census who worked 15-29 hours and had weeks worked of 14-26 or 27-39 were reallocated to full-year work. The net adjustments made to the 1940 numbers were then made by industry and gender to the New Brunswick numbers (the 1950 U.S. census is used for the 1950 New Brunswick adjustments).

B. Relative Income

Computation of relative income begins with deriving weekly wages. For Maine, weekly wages are based entirely on census micro-data. For 1940, 1950, 1980 and 1990 the respondent's weekly wage is annual income from wages and salary divided by the number of weeks worked. For 1960 and 1970, a continuous weeks-worked variable was not available in the public use microdata files, and thus the weekly wage equals wage and salary income divided by the midpoint of the weeks worked category. The categories available in the 1960 and 1970 censuses (and midpoint value adopted) are 1-13 (6), 14-26 (20), 27-39 (33), 40-47 (43), 48-49 (49), and 50-52 (51). For New Brunswick, weekly wages are based on published census data for 1941-61, and census micro-data for 1971, 1981 and 1991 (a categorical weeks worked variable is available in the 1971 census – see section above – and a continuous weeks worked variable is available in both 1981 and 1991; weekly wages are calculated in the same manner as above for Maine). For 1951, weekly earnings itself was not available and so categorical earnings information (total of seven categories, we use the midpoints) are used to compute a weighted weekly earnings average.

For the cell-level analysis, relative income under the 20 weeks worked assumption is then calculated as: $\text{relative income} = [(\text{weekly wages}\{g,y,i,c\} * 20) + \text{UI}\{y,c\}] / (\text{weekly wages}\{g,y,i,c\} * 52)$ where g is gender, y is census year, i is industry and c is country. UI is income from unemployment insurance, and equals the weekly benefit amount times benefit duration. UI for Maine is calculated based on the weekly wage (either the individual's or the cell value) and prevailing UI rules. Benefit durations in Maine are discussed in the text of the paper. For New Brunswick, UI is calculated based on the prevailing UI rules

for a given weeks worked assumption (20 or 30 weeks) with the weekly benefit amount computed based on the schedules available in the on-line appendices, or as 60%/67% (without/with dependents) for post-1971 UI Act census years. In the cell-level analysis (where the presence of dependents is not observed), a 50 percent probability of having dependents is assigned when that affects benefits. For benefit durations for New Brunswick, see Appendix 1.

As Appendix 1 made very clear, weeks of UI eligibility in New Brunswick sometimes depend not just on the number of work weeks in the past year, but on weeks in earlier years as well. In those (relatively rare) cases, we compute eligibility under the assumption the worker has the same annual weeks-worked pattern every year. In addition, a person's UI eligibility sometimes depends not just on the number of weeks worked in a year, but on their precise distribution over time. This is also rare, but whenever the precise distribution matters, we calculate UI benefits for the distribution of weeks that maximizes a worker's UI benefits.

Also, as noted in the paper, our imputed benefit levels focus on an individual who works part-year every year. Under the 1955 UI Act, some "chronic" part-year workers would never qualify for regular benefits, but would qualify for seasonal benefits. To incorporate this policy feature, when regular benefits under the 1955 UI Act enter our analysis, we compute the average number of benefit weeks that the individual would receive over the 1942 - X period (where X is the census year in question) assuming they worked a certain number of weeks every year.

C. Education

For education, the Canadian censuses are the richer source and thus the U.S. variable is used as the base. In general, prior to 1990, the U.S. census tends to report years of education while the Canadian census has a combination of years of education and credential information. The only education variable available in the 1970 and 1980 U.S. censuses used in the micro-data analysis is the 'highest grade of school' variable, which has the following categories: a category for each grade from kindergarten through grade 12 where grade 12 includes both individuals who graduated with a high school diploma and those who did not; and categories for each year of university completed (to a maximum of eight). The 1990 U.S. census changed the education question to a highest level of attainment (i.e., credential). As such, education in the 1990 census is not comparable to previous census years. IPUMS provides a recoded variable, which we use, where they place grade 12 (with and without high school diploma) in a single category, equate 1 to 3 years of college with all college attendance without degree as well as degree below the Bachelor level (e.g., "associate degrees), and then equate 4+ plus years of college with a Bachelor's or greater.

Creating education variables from the Canadian census that are as consistent as possible with the U.S. is done as follows. For 1971, we use grade 11 and under for "less than high school" and grade 12 plus grade 13 as "high school" (includes those with and without a high school diploma, consistent with the US recode); university 1 to 2 years plus university 3 to 4 years without degree completion as "some post-secondary"; university 3 to 4 years with degree, and university 5+ years (with and without degree categories) as "degree". For 1981 and 1991, we use two education variables - 'highest level of school' and 'highest level of elementary and secondary' - in order to create a "high school" variable that includes both high school graduates and grade 12/13 non-graduates. We also include non-university certificates in one of the high-school categories depending on their highest level of elementary/secondary since these are not comparable to college degrees below Bachelor's level in the US such as associate degrees. "Some-post secondary" includes university without a degree and university with certificate/diploma while "degree" includes any university graduate. The only difference in the 'highest level of schooling' variable between the 1981 and 1991 Canadian censuses is that university graduates are separated into BAs, masters and doctorates in the 1991 census.

D. Basic Demographics

Age, gender, marital status and school enrollment are virtually all consistent across years within country and across country. Common-law couples are considered as married in both countries for the 1970 to 1990 census years. For our “children in household” variable, we use the ‘number of children in household’ variable for the U.S. A ‘number of children in household’ variable does exist in the Canadian census, but was only asked of women (and only married or once married women prior to the 1991 census). Thus, for the 1981 and 1991 Canadian censuses we use household type, which indicates the presence of never married children in the household, and should yield a variable that is consistent with the U.S. variable. For 1971, we are forced to use a combination of marital status, size of census family, and – for women only – the children in household variable. For men, we define children as equaling one if family size is greater than two and the individual is married.

E. Part-time work

In the Canadian census, a question is asked, from 1981 through 1991, about whether the weeks worked in the previous year were full-time weeks or part-time weeks. In 1971 ‘usual hours worked’ is available, which asks the individual their usual hours worked from the previous year and should be fairly comparable to the full-time/part-time weeks worked variables in 1981 and 1991. One issue, however, is that the part-time weeks worked question in 1981 did not specifically attach a number of hours to part-time status whereas the 1991 census attached a less than 30 hours per week value. For the U.S. census, ‘usual hours worked’ – based on the previous year – was first asked in 1980, while ‘hours worked last week’ (i.e., hours worked in the census reference week) is consistently available. We use usual hours worked for the 1980 and 1990 U.S. census, but are forced to use hours worked in the reference week for the 1970 U.S. census. Regardless of which hours variable is used, part-time work is defined as less than 30 hours per week in all cases.

F. Industry

We are forced to use the 1971-1991 Canadian census industry variables as the base. Some improvements were made to the industry variables in 1981 and 1991, but for consistency across time we must use the 1971 industry definition which was made available in all three Canadian micro-data censuses. In both countries, industry refers to the individual’s main job (based on hours worked) and is only asked of those working since January 1 of the year previous to the census year (e.g., January 1 1980 in the case of the 1981 Canadian census data). In both countries, for individuals not working during the census reference week, industry refers to the job of longest duration. The 1971 industry variable has 12 categories available: agriculture, forestry, fishing, mining, manufacturing, construction, transportation/storage/communication/utilities, wholesale/retail trade, finance/insurance/real estate, community/business/personal/recreational services, and public administration in addition to a ‘not specified’ (or not classifiable) category. Given the merging of forestry/fishing/mining (the latter being negligible in both New Brunswick and Maine) we are left with eight categories for both the aggregate data and micro-data analyses. We drop the ‘not specified’ category.

For the U.S., we use the IPUMS ‘Industry, 1950 basis’ recode, which is a highly consistent across years. The only across-time inconsistency in this variable was following the 1970 census where coverage was changed from individuals 14+ who had worked in the previous ten years to individuals 16+ who had worked in the previous five years. A total of 148 categories are available. For Canada, the 1970 industry classification is highly comparable to the industries in the 1960 census.¹ The key inconsistencies arise

¹ Given our focus on part-year work, it is worth noting that it is unclear how teachers’ summers have been treated over time. For instance, in the U.S. census in 1940, school teachers were not to count their summer vacations as

when comparing the U.S. 1950 industry codes with the 1971 Canadian industry codes. Given the rich nature of the U.S. data we are able to create reasonably comparable industries across countries after re-coding 'logging' from manufacturing (in the U.S.) into primary (for Canada). With public-use data, however, there are some inconsistencies that cannot be corrected for. Our reading of the U.S. 1951 classification (see U.S. Bureau of Census, *Alphabetic Index of Occupations and Industries: 1950*, Washington, D.C., 1950) relative to the Canadian 1971 classification is that most inconsistencies come from specialized services specific to an industry. For instance, the U.S. places oil and gas field services in mining while Canada places it in professional services. On the other hand, Canada places research and testing services incidental to agriculture in agriculture while the U.S. places them in professional services. Various examples go both ways, but they are largely limited to highly specialized services, which are unlikely to amount to substantive inconsistencies overall. Manufacturing, trade, finance etc., and the primary sector (after re-coding logging) are highly consistent across the two countries.

G. Wage and salary income

Our measures of wage income are highly consistent, both over time and between countries. The level at which these variables are top-coded however does vary however. In both countries, wage income refers to gross wages and salaries, before deductions (for items such as income tax, pensions, union dues, unemployment insurance), for the previous calendar year (i.e., the year to which the weeks worked variables refer). Payments-in-kind or reimbursements for business expenses are excluded in both countries. Commissions, tips, bonuses, piece-rate payments, and military pay are included in both countries (military bonuses were excluded in the U.S. for 1950 – 1970 censuses).

time worked, but in 1950 teachers were to report all weeks paid (including paid absences, vacations, and sick time). In later years, there were no specific instructions, and it would seem that teachers who received pay over the summer would have considered this to be paid vacation and therefore would have included it as weeks worked.