Rural-Urban Migration and Socioeconomic Mobility in Victorian Britain

Jason Long, Colby College
(Comments welcome: jmlong@colby.edu)

I would like to thank my advisors, Joel Mokyr, Joe Ferrie, and Joe Altonji, for their helpful comments. I also benefited from conversations with Henry Siu, Jim Sullivan and Chris Taber, and from input from workshop participants at Northwestern University and participants of the 2002 All-UC Group in Economic History Conference and the 2002 Cliometrics Conference. Justin Hayes and Humphrey Southall provided me with data. This research was greatly aided by a Northwestern University Graduate Research Grant and a Pew Younger Scholars Program Fellowship.

Victorian Britain experienced the most rapid and thorough urbanization the world had yet seen. People left the countryside for the cities to such an extent that, by mid-century, Great Britain had become the first predominantly urban society in history. This study addresses two related questions: Who were the rural-urban migrants, and what were the returns to rural-urban migration? Specifically, I examine the nature of the selection mechanism by which some moved to urban areas and other remained rural, and I assess the treatment effect of migration in terms of attaining higher socioeconomic status. I answer these questions with the aid of a new dataset of 28,000 individuals matched between the 1851 and 1881 Censuses of the Population of England and Wales. I build a structural econometric model to assess the effect of migration and the nature of migrant selection, controlling for the endogeneity of the migration decision. I find that those who left the countryside for the cities did not come from the bottom of the economic and social ladder; rather, they were better off initially than those who remained in rural areas. Urban migrants were positively selected: their prospects in the urban labor market were superior to those who did not move. Moreover, their prospects in the rural labor market were also superior. Finally, the decision to move was, by and large, a fruitful one. The treatment effect of urban migration was large and positive across all socioeconomic strata. On average, people who moved to the city were substantially more successful in improving their socioeconomic status than they would have been had they remained in rural areas, and they were more likely to experience upward intergenerational occupational mobility.

Unlike the early 20th century, which saw the rise of national social welfare programs, virtually no adverse institutional incentives existed at the time to inhibit the movement of Britain's people. And move they did; high rates of internal migration, and rural-urban migration in particular, represented one of the most prevalent demographic features of 19th century Britain. This mobility also had important efficiency implications for Britain's growing economy, as labor reallocated itself from areas of surplus to areas of scarcity. It is understandable that the topic of

internal migration in 19th century Britain has received extensive attention in the literature. A recent survey lists 15 books and articles from the pioneering work of Ravenstein in 1885 to the present, covering every decade from the 1920s to the 1990s.[1] Each of these studies aims to describe and explain patterns of internal mobility, generally at some aggregate level such as the county. None of them actually observes individuals making the migration decision.[2] Much is learned with these aggregate methods, but much remains unknown. The reason for the omission is simple: until now, no nationally representative micro-level panel data have been available with which to observe the geographic and occupational mobility of individuals over time. Such data were constructed for the present study, which poses some new questions, chiefly—Who were the rural-urban migrants, and to what extent did they benefit from the move?

The Data

The data used here come from a new sample of approximately 28,000 males linked from the 1851 Census of the Population to the 1881 census. The population censuses have long been recognized as the most important source of individual-level data for Great Britain from the 19th century. They are large, and they represent essentially the entire population. Each census is only a cross-sectional snapshot of the population, however. Successive censuses can be used to examine changes in the nation over time but not changes in the lives of individuals over time. No continuity exists for individuals between the censuses. But continuity can be created. The Genealogical Society of Utah in conjunction with the Federation of Family History Societies has recently computerized the entire 1881 census of the population of England, Wales, and Scotland. With these data any individual or group of individuals can quickly and easily be searched for throughout the entire 1881 census. In this case the search was for individuals from another census: that of 1851. The dataset used was a computerized 2 percent sample of the 1851 census, compiled principally by Anderson, Collins, and Stott. [3]

The information available for each individual comes from the questions asked in the two censuses: name, address, relationship to head of household, marital status, age, sex, occupation, county and parish of birth, and whether blind or deaf and dumb. Name, age, and birthplace information were used to link individuals between the two censuses. In order to be considered a

[1] Boyer and Hatton, "Migration." Notable are Ravenstein, "Laws," Redford, *Labour Migration*, and Baines, *Migration*.
[2] Williamson, *Coping*, includes micro-level analysis, but on individuals observed only at a point in time. Pooley and Turnbull, *Migration and Mobility*, examine panel data on individuals, but their data are chronologically diffuse, it is of questionable representativeness, and they do not include any econometric analyses of migration.
[3] Both datasets available from the Data Archive at the University of Essex: 1881 census as study #3643, 1851 sample as #1316.

true match for an individual from 1851, an individual from 1881 had to have the same name or a close phonetic variation thereof, a year of birth different by no more than five years, and the same county and parish of birth. The variation in birth year was allowed in order to account for age misreporting, and the parish of birth, like the name, was allowed to vary. If by these criteria an individual from the 1851 sample had more than one match from the 1881 census, then that individual was discarded.

Applying this matching process to an initial pool of 187,117 English and Welsh males from the 1851 2 percent sample yielded a set of 28,474 men observed both in 1851 and 1881, a success rate of approximately 15 percent. The data come from two nationally representative sources, so as long as the matching process does not skew the sample, the set of matched individuals should also be representative of the population of England and Wales in 1851.[4] Table 1 shows a comparison between the sample of matched individuals and the entire group of males from the 1851 2 percent sample.[5] There are differences between the groups throughout the categories, but they all stem from one important factor: the matched men are younger on average. The average age for the matched men is 18, while for the entire group it is 24. This should be expected. Life expectancy at birth was only 39.5 years in England in 1851; younger men were more likely still to be alive in 1881. From this one difference follow the others: matched men were more likely to be sons rather than household heads; they were less likely to be married; more likely to be students and weavers and less likely to be anything else; and more likely to still be living within their county of birth (having had less time to move away).

In order to examine the rural-urban migration decision, a smaller subsample must be extracted from the 28,000 matched individuals. First, in order to see what caused some to move to the city and others to remain behind, and what happens to those who go compared to those who stay, it is necessary to consider only individuals who began the period outside of any city.[6]

Second, attention will be restricted to those who were listed as sons in the 1851 Census. This group represents the largest in the matched sample, at 60 percent. (25% were household

[4] One important feature of the matching process is that only survivors over the 30-year period are included in the sample. In 19th century Britain, urban areas exhibited higher mortality rates than rural. To the extent that this urban mortality bias was more severe for the poor, estimates of the returns to migration will be biased upward.

[5] It also includes the smaller subsample of individuals to be used in the econometric analysis. This group is discussed below.

[6] There is no hard definition for what constitutes a "town" or a "city." The UN has recommended that all places with more than 20,000 inhabitants living close together be considered as "urban." The U.S. Census, uses 2,500. These two numbers are often considered to be meaningful cut-off points. In 19th century Britain, the cities of 20,000+ were experiencing the most rapid growth, and it will be these that constitute "urban" for the purposes of this study. The populations of towns and cities in both1851 and 1881 were calculated using the 1881 census to ensure that the effects of increasing population between the two years did not cause people to appear to urbanize simply by living in a growing town.

TABLE 1
COMPARISON OF ALL MALES, MATCHED SAMPLE, AND ESTIMATION SUBSAMPLE

	All males (%)	Matched sample (%)	Estimation subsample (%)
Age (mean)	24.36	18.09	16.02
Occupation			
Student	15.84	22.30	27.61
Ag. Laborer	11.22	11.96	18.76
Farmer	3.01	2.20	0.45
Laborer	2.91	2.30	2.81
Miner	2.58	2.11	3.05
Weaver	2.11	2.66	3.26
Tailor	1.68	1.59	1.51
Carpenter	1.61	1.68	1.80
None	1.29	1.68	0
Shoemaker	1.12	1.27	1.75
Relation			
Son	44.56	59.86	100
Head	33.80	25.24	0
Lodger	5.69	3.46	0
Servant	1.89	3.69	0
Visitor	1.86	1.32	0
Marital status			
Married	47.72	42.31	4.12
Unmarried	47.66	55.61	95.21
Widowed	4.18	1.45	0.22
Region			
East	6.58	7.88	9.86
Lancashire-Cheshire	11.31	12.43	8.77
London	8.55	5.14	0
London Environs	12.70	13.71	13.99
Midlands	19.89	20.94	25.12
North	5.06	5.02	5.25
South	18.49	20.57	26.50
Wales	4.93	4.61	3.13
York	12.50	9.70	7.39
Lived in county of birth in 1851	71.56	88.41	95.12
N	187,117	28,474	3,774

Notes: All values are from 1851. The occupations listed are the 10 most prevalent among the sample of all males. Students were children either attending school or receiving formal instruction at home. *Sources*: 1851 census 2% sample and new sample of matched individuals.

heads; 15% were grandsons, visitors, etc.) More importantly, for the sons, not only can we observe their own occupation, but also that of their father, allowing us to gauge the impact of rural-urban migration on both intra- and intergenerational socioeconomic mobility. Furthermore, the location choice of sons in 1851 is likely to be exogenously chosen by the head of household; including household heads, who would already have made a rational 1851 location choice, would introduce endogeneity into the problem.[7]

Finally, in order to judge the effects of moving, or not moving, to the city, the sample must also be limited to those individuals for whom there is solid economic information from both censuses. Unlike their U.S. counterparts, the Victorian censuses do not provide any

[7] Only sons aged 9 (the minimum legal working age) to 29 in 1851 were included.

quantitative economic variables such as personal or real estate wealth. What they do include is fairly detailed occupational information. A distinction was made between employer and employee and master and apprentice, and the number of persons employed was recorded. There were literally thousands of occupations listed in the enumerators' books; there are in the neighborhood of 8,000 unique occupations given in the set of 28,000 matched individuals. In the absence of wage information, some system of ranking jobs according to their desirability must be employed. Fortunately, fairly strong consensus has arisen in support of a ranking scheme proposed by Armstrong. He argues that only the Registrar General's social classification schemes of 1921 and 1951 satisfy the dual requirements of being appropriate for the Victorian census data and at the same time including exhaustive published lists for categorizing the vast array of occupations listed in the census. Armstrong's scheme consists of five ranked classes of occupation: I – Professional, II – Intermediate, III – Skilled, IV – Semiskilled, and V – Unskilled.

This system of classification is clearly aligned to capture job desirability. It makes a normative statement: higher-class jobs were better than lower class. Two sources, both published by the General Register Office of Great Britain, are needed to classify each occupation fully. The *Classification of Occupations*, 1921, lists approximately 16,000 different occupations and gives a three-digit code number for each. The *Decennial Supplement* for 1921 then provides the appropriate class rank for every code number.[8] Armstrong provides several modifications to this system to bring it into somewhat better harmony with the nature of nineteenth century occupational and class structure. Probably most important is his use of the employee information given under the occupational field in the census. Regardless of job title, all employers of 25 or more are to be placed in Class I, and all people with Class III or IV occupations employing at least one person other than a family member are to be placed in Class II. Armstrong's basic scheme has been modified in one way, in order to take full advantage of the information offered by the census. Since entire households are observed both in 1851 and 1881, the ratio of servants to household members can be calculated for each household. The job class ranking of some individuals was upgraded (never downgraded) according to the following scheme, proposed by Royle: all heads whose households contained at least one servant per household member were placed in Class I, all others with one servant per three household members in Class II, and any

[8] Thanks to Dr. Humphrey Southall for his computerized dictionary with which I was able to code some of the occupations.

others that employed at least one servant in Class III. Formulated in this way, the job class variable may be thought of as a representation of socioeconomic status, and the terms occupational class and socioeconomic status are used interchangeably throughout the study.

Out of the 28,474 matched individuals, 3,774 are sons living with their father in a rural area in 1851, with both son and father reporting solid occupational information. They are shown in the third column of Table 1 for comparative purposes, though of course this subsample is unrepresentative of the male population of England and Wales by construction. It is this group which will be used to analyze rural-urban migration.

A Model of Rural-Urban Migration

Migration to the city is not a randomly assigned treatment; an individual's expectation of his labor market outcome will influence his decision of whether or not to move. The model used to analyze the migration decision, then, is one of regime switching with endogenous switching. The basic model is

$$\text{Regime 0:} \quad y_{0i} = \beta_0' X_{0i} + \varepsilon_{0i} \quad \text{if } D_i = 0 \tag{1}$$

$$\text{Regime 1:} \quad y_{1i} = \beta_1' X_{1i} + \varepsilon_{1i} \quad \text{if } D_i = 1 \tag{2}$$

$$D_i = 1 \quad \text{if } \gamma' Z_i + u_i \geq 0 \tag{3}$$

$$D_i = 0 \quad \text{otherwise}$$

where y represents the labor market outcome and D the decision of the regime in which to participate. X and Z are the factors that influence an individual's labor market outcome and participation decision, respectively, β and γ are vectors of coefficients, and ε and u are unobservable factors. A wide range of empirical questions has been examined with this model, typically in the form of switching regressions, where y is a continuous variable, often wage, and equations (1) and (2) are estimated by OLS. The switching regression model cannot be applied in the present setting, because here the labor market outcome variable—socioeconomic status—is discrete, taking one of five ranked values. The ordinal nature of the socioeconomic status variable suggests the use of ordered probit analysis. In place of the standard model, a model of switching ordered probits is developed.

The model is defined as follows. There exists a continuum of job quality, $Y^* \in (-\infty, +\infty)$. Individuals have a utility function $\mathrm{U}(y^*, \Theta)$, where y^* is a realization of Y^* and Θ is a vector of other inputs. All individuals prefer higher quality jobs: $\partial \mathrm{U} / \partial y^* > 0$. The maximum quality of job that an individual can attain in 1881, y_i^*, is defined to be a linear function of that person's

observable traits and skills, X_i , and unobservable characteristics, ε_i . The function may be different in the urban regime than in the rural, though the relevant elements of X_i are assumed to be the same in each:

$$\text{Rural:} \qquad y_{0i}^* = \beta_0' X_i + \varepsilon_{0i} \tag{4}$$

$$\text{Urban:} \qquad y_{1i}^* = \beta_1' X_i + \varepsilon_{1i} \tag{5}$$

The decision facing each individual in this model is whether to migrate to a city or remain in some rural setting. The net benefit of moving, D_i^* , is defined to be a linear function of (1) Z_i , a vector of observable individual and location specific characteristics, (2) the difference between the individual's maximum attainable job quality in the urban regime versus that in the rural, $\left(y_{1i}^* - y_{0i}^*\right)$, and (3) u_i , unobservable characteristics:

$$D_i^* = \gamma_1' Z_i + \gamma_2 \left(y_{1i}^* - y_{0i}^*\right) + u_i \tag{6}$$

Only the outcome of the choice, D_i , is observed: $D_i = 1$ if $D_i^* \geq 0$, $D_i = 0$ otherwise. To this point, the model is identical to the base model specified in (1) – (3). The difference is that in the base model, either y_{0i} or y_{1i} is observed for each individual, depending on the value of D_i , whereas in the current model neither y_{0i}^* nor y_{1i}^* is observed for any individual. What is observed is the job class, y_i , of every individual, either in the rural or urban regime:

$$\begin{aligned} y_{0i} &\in [1,2,3,4,5] \qquad \text{if } D_i = 0 \\ y_{1i} &\in [1,2,3,4,5] \qquad \text{if } D_i = 1, \end{aligned} \tag{7}$$

The following relationship is assumed to exist between job class, y_i, and job quality, y_i^* :

$$\begin{aligned} y_{ri} = \quad & 5 \quad \text{if} \quad -\infty < y_{ri}^* \leq k_1 \\ & 4 \quad \text{if} \quad k_1 < y_{ri}^* \leq k_2 \\ & 3 \quad \text{if} \quad k_2 < y_{ri}^* \leq k_3 \\ & 2 \quad \text{if} \quad k_3 < y_{ri}^* \leq k_4 \\ & 1 \quad \text{if} \quad k_4 < y_{ri}^* \leq +\infty \qquad r = 0,1 \end{aligned} \tag{8}$$

where $k_1 - k_4$ are constants. Class I jobs are the most desirable, and Class V the least. Under this formulation, all jobs within a class are not equivalent; indeed, the best job in any class is only marginally inferior to the worst job in the next highest class. Also, the quality/class structure is identical in the rural and urban labor markets: the four threshold levels $k_1 - k_4$ do not vary between the two regimes. If this were not the case, an individual could change job class without

changing jobs simply by moving from one regime to the other.

The three-equation system defined by (6) and (8) is estimated by Full Information Maximum Likelihood (FIML), which proceeds in two stages. First, by substituting (4) and (5) into equation (6), (6) can be written in reduced form, in terms of observables, as

$$\begin{aligned} D_i &= 1 \qquad \text{if } \gamma' W_i + v_i \geq 0 \\ D_i &= 0 \qquad \text{otherwise} \end{aligned} \tag{9}$$

where W contains all the elements of X and Z. A standard assumption is made regarding the disturbance terms: ε_0 , ε_2 , and v are assumed to be i.i.d. draws from a trivariate normal distribution with mean vector zero.

With this distributional assumption, (8) becomes a standard ordered probit model with five outcomes and four threshold levels, with the outcome depending on the latent variable y^*. The three equations in (8) and (9) are jointly estimated by maximizing the likelihood function:

$$\mathrm{L}(\beta_0,\beta_1,\gamma,k_1,k_2,k_3,k_4,\rho_0,\rho_1) =$$

$$\left[\prod_{y_i=5}\mathrm{F}(k_1-\beta_0'X_{0i},\gamma'W,-\rho_0)\prod_{\substack{y_i=j\\ j=2,3,4}}\left[\mathrm{F}(k_j-\beta_0'X_{0i},\gamma'W,-\rho_0)-\mathrm{F}(k_{j-1}-\beta_0'X_{0i},\gamma'W,-\rho_0)\right]\prod_{y_i=1}\mathrm{F}(\beta_0'X_{0i}-k_4,\gamma'W,\rho_0)\right]^{1-D_i}$$

$$\left[\prod_{y_i=5}\mathrm{G}(k_1-\beta_1'X_{1i},-\gamma'W,\rho_1)\prod_{\substack{y_i=j\\ j=2,3,4}}\left[\mathrm{G}(k_j-\beta_1'X_{1i},-\gamma'W,\rho_1)-\mathrm{G}(k_{j-1}-\beta_1'X_{1i},-\gamma'W,\rho_1)\right]\prod_{y_i=1}\mathrm{G}(\beta_1'X_{1i}-k_4,-\gamma'W,-\rho_1)\right]^{D_i}$$

where F and G are, respectively, the bivariate normal distribution functions of (ε_{0i} , v_i) and (ε_{1i} , v_i) and ρ_0 and ρ_1 are the correlation coefficients of the two distributions. β_0 and β_1 are assumed to include a constant term, so the four threshold levels, k, are not all identified, and k_1 is normalized to 0. This estimation procedure produces consistent and asymptotically efficient estimates of the structural parameters of interest, β_0 and β_1 .

The second stage is the recovery of the structural parameters of equation (6). With estimates of β_0 and β_1 in hand, predicted values of both y_{0i}^* and y_{1i}^* are obtained for each individual as $\hat{y}_{ri}^* = \hat{\beta}_r' X_i$, r = 0,1. These predicted values are substituted into (6) in place of the unobservable y_{0i}^* and y_{1i}^* . (6) is then estimated by probit maximum likelihood to obtain estimates of the structural parameters γ_1 and γ_2.

The estimates of β_0 , β_1, γ_1, and γ_2 reveal the determinants of job class attainment and

urban migration. To analyze the selection of urban migrants and rural persisters and the treatment effect of migration, it is necessary to define several more parameters to be estimated. The selection of urban migrants (s_1) and the selection of rural persisters (s_0) are given by

$$\begin{aligned} s_1 &= \mathrm{E}\left(y_1^* \mid D=1\right) - \mathrm{E}\left(y_1^* \mid D=0\right) = \widetilde{X}_1\hat{\beta}_1 - \widetilde{X}_0\hat{\beta}_1 \\ s_0 &= \mathrm{E}\left(y_0^* \mid D=0\right) - \mathrm{E}\left(y_0^* \mid D=1\right) = \widetilde{X}_0\hat{\beta}_0 - \widetilde{X}_1\hat{\beta}_0, \quad \text{where } \widetilde{X}_r = \mathrm{E}\left(X \mid D=r\right), r=0,1 \end{aligned} \tag{10}$$

If $s_1 > 0$ then, conditional on the observables, urban migrants were positively selected; i.e., they achieved higher job quality in the urban labor market than the rural persisters would have had they chosen to move to the city. The analogous characterization holds for s_0 and the selection of rural persisters.

The treatment effect of rural-urban migration, τ, is defined to be

$$\tau = \mathrm{E}\left(y_1^* - y_0^* \mid D=1\right) = \widetilde{X}_1\hat{\beta}_1 - \widetilde{X}_1\hat{\beta}_0 \tag{11}$$

Like the selection parameters, τ is calculated conditional on the observables of the model. The treatment effect is considered in terms of the average urban migrant: $\widetilde{X}_1$ represents the endowments of the skills and attributes available to the average migrant. If the treatment effect is positive, $\tau > 0$, then migration to the city yielded a positive return in terms of job quality for the average rural individual in England and Wales of the type who typically chose to move to the city from the countryside.[9]

Estimation of the Model

The determinants of maximum attainable job quality in 1881—the elements of X in (4) and (5)—include class rank in 1851, father's class rank, age, age squared, age discrepancy, whether the individual lived in a town in 1851, whether the individual was an eldest son, whether the father was a farmer or employer, an interaction between the previous two terms, the industrial classification of the individual's occupation in 1851, the degree of age-heaping for the county of residence in 1851, and region in 1851. Summary statistics for these variables, and for those in the migration equation, are presented in Table 2. Several of the variables merit some explanation. The inclusion of the individual's class rank in 1851 allows us to measure the effect of moving to

[9] The matched data measure only permanent moves; $D_i = 1$ if individual i moved to an urban area after 1851 and remained there until 1881. τ therefore measure the treatment effect of permanent urban migration. Some of course will move to the city and subsequently decide to return, quite possibly because of a poor labor market outcome. Such individuals are not counted as movers in the present analysis. Greater frequency of observation would permit a refinement of the analysis, but the exclusion of at least some return migrants is an inherent feature of empirical migration studies.

the city on the change in an individual's job class rank; it controls for individual-specific effects on the level of job quality. This is analogous to estimating a difference equation without constraining the coefficient on the first time period value to be equal to one. The model is formulated this way in order to estimate the effect of class in 1851 on class in 1881 as an indicator of the rigidity of occupational class structure in Victorian Britain, and in particular to estimate the difference of that effect between those who switch regimes and those who do not. The job class of the father is also included. Just as estimating the effect of the individual's class in 1851 on his class in 1881 allows us to measure the degree of occupational mobility, so including father's class in 1851 reveals the degree of intergenerational mobility, an important question itself. So the model is really one of change in job class, both intra- and intergenerational.

One class included in 1851 is not a proper occupational classification at all: that of "student." It is the second most common class in 1851, and the single most commonly reported occupation, in this sample as well as in the census as a whole. It does not fit into the Registrar General's occupational classification system, but it may be regarded as a class of its own. For the 1851 census, parents were instructed to report their children as students ("scholars" in the language of the Victorian censuses) if they were older than five and were "daily attending school, or receiving regular tuition under a master or governess at home." Though no substantive details of their instruction can be known, children older than nine who were listed as students were receiving an education. Many of their peers were not; they were working. By including the students as a separate class, we can gauge in a very rough fashion the effect of early schooling on future job quality attainment as compared to employment at a young age.

Literacy and specific educational information are not observed, but a proxy can be constructed. "Age discrepancy" is defined to be

$$\left|\text{Reported Age}_{1881} - \left(\text{Reported Age}_{1851} + 30\right)\right|$$

As discussed above, the age matching criterion is that an individual's reported age in 1881 cannot be more than five years different from what it should be (the reported age in 1851 plus 30). Roughly half of all individuals in the sample reported consistent values for age in the two censuses. Another quarter was off only by a year, and the remaining quarter was off by two to five years. This was a time before systematic record keeping, and many people had only an approximate idea of their age. The observed age discrepancy for each individual gives an

TABLE 2
SUMMARY STATISTICS, ESTIMATION SAMPLE

		Mean, 1851			Mean, 1881		
		All	Urban	Rural	All	Urban	Rural
INDIVIDUAL CHARACTERISTICS:							
Job class:	1 - Professional	0.004	0.01	0.003	0.02	0.03	0.02
	2 - Intermediate	0.02	0.02	0.02	0.15	0.14	0.16
	3 - Skilled	0.30	0.32	0.30	0.45	0.56	0.42
	4 - Partly skilled	0.28	0.20	0.30	0.25	0.12	0.29
	5 - Unskilled	0.12	0.10	0.12	0.12	0.15	0.11
	Student	0.28	0.36	0.26		---	---
Father's class:							
	1 - Professional	0.02	0.03	0.01			
	2 - Intermediate	0.16	0.17	0.16			
	3 - Skilled	0.38	0.45	0.36			
	4 - Partly skilled	0.33	0.27	0.35			
	5 - Unskilled	0.10	0.08	0.11			
Age discrepancy (years):	0				0.49	0.49	0.49
	1				0.29	0.28	0.29
	2				0.10	0.08	0.11
	3				0.05	0.07	0.05
	4				0.03	0.04	0.03
	5				0.03	0.04	0.03
Age (years)		16.02	15.26	16.26			
Eldest		0.57	0.57	0.58			
Inheritance		0.12	0.11	0.12			
Married		0.02	0.02	0.03			
Not in town of birth		0.26	0.32	0.24			
Industry:	Agriculture	0.24	0.15	0.27	0.24	0.04	0.30
	Building	0.03	0.03	0.03	0.08	0.10	0.07
	Distributive	0.02	0.02	0.02	0.09	0.13	0.08
	Mining	0.04	0.02	0.04	0.05	0.03	0.05
	Textiles	0.08	0.12	0.06	0.04	0.06	0.04
	Iron & Steel	0.02	0.03	0.02	0.04	0.06	0.03
	Other Manuf.	0.15	0.15	0.15	0.20	0.23	0.19
	Other	0.42	0.48	0.40	0.26	0.36	0.23
LOCATION CHARACTERISTICS:							
Living in city		0	0	0	0.24	1	0
Living in town		0.51	0.63	0.47	0.40	0	0.53
Region:	Lancashire	0.06	0.11	0.04	0.08	0.20	0.04
	London	0	0.00	0.00	0.04	0.17	0.00
	London Environs	0.14	0.15	0.14	0.14	0.14	0.14
	Wales	0.03	0.02	0.04	0.03	0.02	0.04
	Yorkshire	0.07	0.08	0.07	0.08	0.09	0.07
	Others	0.70	0.64	0.72	0.63	0.39	0.71
Age-heaping (index)		61.00	61.58	60.82			
Distance to city (km)		23.35	19.60	24.51			
Wage gap (sh/wk)		4.79	4.98	4.73			
Previous migrants		13,716	15,339	13,211			
Unemployment (percent)		7.97	8.84	7.70			
Nearby cities		32.56	36.77	31.25			
Pct. in agriculture (percent)		29.29	0.26	0.30			
Pct. in manuf. (percent)		23.00	0.26	0.23			
N		3,774	896	2,878	3,774	896	2,878

Notes: Variable definitions and sources given in text. "All" refers to all individuals in the estimation subsample. "Urban" refers to those who moved to an urban place between 1851 and 1881, "Rural" to those who remained in any rural place.

indication of that individual's familiarity with arithmetic ("numeracy") and the precision of his concept of time. In general it may be regarded as a proxy for literacy, education, intelligence, and the like, and therefore be expected to have a positive impact on job quality attainment.[10] For the population as a whole, people not knowing their age precisely manifests itself as age heaping—the tendency of any population to over-report rounded ages. The age heaping variable is a county-level index of this tendency and is intended to capture otherwise unmeasured community effects that might influence the labor market performance of individuals.[11]

The remaining variables are more straightforward. People living in a town in 1851, as opposed to those in the truly rural countryside, may have possessed a form of locational human capital that would ease their transition into the big-city labor market. This effect is captured by the inclusion of a dummy variable that equals one if the individual lived in a town of between 2,500 and 19,999 inhabitants. The dummy variables indicating whether a person in the sample was an eldest son and whether his father was a farmer or an employer are designed to capture an inheritance effect. Fathers who owned land were likely to leave it to their eldest son, as the system of primogeniture was the norm with regard to property inheritance in most of England. Likewise, sons of fathers who owned any sort of business were likely heirs to an occupational inheritance, though in this case the inheritor well could have been any son rather than just the eldest. Finally, industrial and regional dummy variables are included, the former being classified according to the well-known scheme developed by Booth.

In addition to his expectation of his future job prospects in both regimes, the factors that influence an individual's decision whether to move to the city—the components of Z in (6)—include both individual- and location-specific elements. The individual-specific characteristics include age and age squared, marital status, age discrepancy, whether the person lived in a town in 1851, whether the person was living in the parish of his birth in 1851, and industrial classification. The human capital interpretation of migration (that people migrate in order to maximize lifetime net benefit from moving) would suggest that the likelihood of migrating will decrease with age, as individuals have a shorter time span over which to reap the gains from moving and as they make more location-specific investments in their place of residence. More educated, knowledgeable people would have a lower tendency to report their age inconsistently;

[10] There is, in fact, a correlation between age discrepancy and being a student in 1851: students were 8 percent more likely to know their exact age than were non-students.

[11] For a more thorough discussion of the economic usefulness of age heaping information, particularly in a migration analysis setting, see Mokyr, *Ireland*.

these same people should be better able to gather information on potential urban moves. Moving to a big city might have been a less drastic change, and so have carried a lower psychic cost, for those already living in towns in 1851. There are only two new variables in this group. The first is marital status, the expected effect of which is not clear *a priori*. While studies of overseas migration generally find the typical migrant to be a young, single male, this may not be the case for shorter-distance, often local moves from countryside to city. Indeed, during the second half of the nineteenth century, the decline in rural employment affected the job prospects of women more than those of men, so women became more likely than men to leave the countryside in order to find work in the city, most often as domestic servants.[12] The second new variable is an indicator of whether an individual was living in his parish of birth in 1851; such individuals might be expected to have closer ties to their community and be less likely to move.

The elements of Z that depend on the individual's location in 1851 include the distance to the nearest city, the number of large cities within 100 km, a proxy for the total number of previous migrants in each nearby city, the average rural-urban wage gap, the urban unemployment rate, an interaction between the unemployment rate and distance to nearest city, the percent of the male workforce engaged in agriculture and in manufacturing in the home county in 1851, the degree of age heaping of the home county in 1851, and the regional dummies.[13] Of these variables only age heaping and the regional dummies are in X. Distance to the nearest city and number of large cities nearby were calculated using modern grid point references for place names from the 1991 British Census.[14] Distance should be negatively correlated with tendency to migrate, as it would be both more costly to move to a distant city and more difficult to gather pertinent information. Knowing people who had already moved to a particular city was another important source of information for the potential migrant. A proxy was constructed for this friends-and-family effect. First, the county of birth of all urban residents in England and Wales in 1881 was tabulated by city using the 1881 census data. Second, the set of most likely destination cities for each individual in the sample was defined to be the nine closest cities to his parish of residence in 1851 plus London. Finally, for each individual i with likely destination cities $j = 1,\ldots,10$, the proxy for the stock of previous migrants is defined to be

[12] Baines, "Population."
[13] Share in agriculture and manufacturing are from 1871 census, in Lee, *British Regional Employment Statistics*.
[14] Thanks to Justin Hayes with MIMAS at the University of Manchester for providing me with these data.

$$previous\ migrants_i = \sum_{j=1}^{10} \left(\frac{d_{ij}^{-1}}{\sum_j d_{ij}^{-1}} \right) m_{ij} \tag{12}$$

where m_{ij} is the number of people living in city j in 1881 born in the county of residence of individual i, and d_{ij} is the distance between city j and the place of residence of individual i. This is a proximity-weighted average of the m values of all likely destination cities. The average urban-rural wage gap and the urban unemployment rate were calculated along similar lines. The wage gap is defined as the average wage of laborers in the building trade in nearby cities minus the average wage of agricultural laborers in the origin county.[15] The urban unemployment rate was calculated using data on joblessness among members of the Amalgamated Society of Engineers, available for 56 cities in England and Wales from 1858 through 1909.[16] It is calculated as the proximity-weighted average—as in (12)—of the rates from 1862, 1868, and 1879 for the two nearest cities and London. This is the urban unemployment rate, so its expected effect on tendency to move to the city should be negative; people would be less likely to move to the city if there were high unemployment there. However, we do not observe the relevant local unemployment rate and so cannot control for it. Therefore, since the urban unemployment rate is taken from the cities nearest to each individual's place of origin, it will proxy, to an extent, for the local unemployment rate facing each individual. Insofar as this is true, the effect of unemployment would be positive. To test for this effect, an interaction term between the unemployment rate and distance to nearest city is included. This term is anticipated to have a negative coefficient: for any given unemployment rate, an increase in the value of the interaction term indicates that the unemployment pertains to a city which is farther away, thus more truly the unemployment of the destination city and less that of the place of origin.

The three parameter vectors β_1, β_2, and γ are identified under the assumptions of the model; however, if the variables in X and Z are identical, this holds only if the structure and normality assumptions of the model are exactly correct. Fortunately, in this case there are some reasonable exclusion restrictions. Z includes quite a few variables not present in X: marital status, whether the person was living in the parish of his birth in 1851, and all of the location-specific variables save age heaping and the regional dummies. X does not include quite so many unique

[15] Agricultural laborers' wages are from Hunt, "Industrialization," Table 6. They cover the years 1867-1870. Building laborers' wages are from Hunt, *Regional Wage Variation*, Table 1.5, for the year 1886.

[16] These data are from Southall, "Regional Unemployment," Table 3. For a discussion of its appropriateness as a proxy for the unemployment experiences of other groups, see Southall and Gilbert, "Good Time to Wed."

variables: the eldest and inheritance dummies, father's job class, and own job class in 1851. These, of course, do influence the individual's migration decision, but only through their effect on $\left(y_{1i}^{*} - y_{0i}^{*}\right)$. With these exclusion restrictions, the parameters of the model are identified, even if the assumptions of the model do not hold exactly.

Empirical Results

Occupational Class Attainment

Estimating the two job class equations, (8), yields some general insight on the nature of the attainment of socioeconomic status in 19th century Britain, and on how that process differed between the rural and urban areas. Results from estimating (8) as well as from estimating what may be considered the baseline model—an ordered probit estimation of job class in 1881 on X plus a dummy variable for moving to a city by 1881—are presented in Table 3. The columns labeled "Urban" show the estimates of β_1, the coefficients for those who chose to move to an urban area by 1881, and the columns labeled "Rural" show the estimates of β_0, the coefficients for those who remained in a rural place throughout the time period. The ancillary parameters ρ and k do not vary between the two groups. Interpretation of ordered probit results is not exactly straightforward. The coefficient estimates represent the effect of a change in the explanatory variable on the unobserved latent variable y^{*}, job quality in this case. Job quality is a theoretical construct of the model; it has no units or ready interpretation. The coefficient estimates reveal the sign of the overall effect of the explanatory variable and the precision with which the effect is estimated. For an understanding of the magnitude of the effect on the observed variable, here job class, the marginal effect of each dependent variable must be calculated. These are shown in Table 4, with the "Urban" and "Rural" columns as in Table 3. Also shown are the baseline probabilities of attaining each job class conditional on moving to a city or remaining in a rural place. Each explanatory variable has a different marginal effect for each possible outcome of the ordered probit, so there will be a different effect for each of the five possible job classes that an individual could attain in 1881. The sign of the overall effect of each explanatory variable is unambiguous, but the sign will vary across 1881 job classes. A variable that positively influences job quality will positively influence the probability of being in the highest job class but will negatively influence the probability of being in the lowest job class.

As expected, the 1851 class variables, both own and father's, are strong influences on subsequent job quality attainment. In each case save one, the effects are as anticipated: being in a

TABLE 3

MAXIMUM LIKELIHOOD RESULTS: DETERMINANTS OF JOB CLASS ATTAINMENT

Variable	Class, 1881	Class, 1881		Variable	Class, 1881	Class, 1881	
		Urban	Rural			Urban	Rural
	(1)	(2)	(3)		(1)	(2)	(3)
Constant	0.449	-0.131	0.699	*Industry*	(***)		(***)
	(0.319)	(0.758)	(0.359)	Agriculture	-0.181	-0.199	-0.128
Own class, 1851	(***)	(***)	(***)		(0.106)	(0.234)	(0.122)
1,2 - Prof., Intermed.	1.371	1.958	1.156	Building	0.280	0.305	0.301
	(0.147)	(0.289)	(0.172)		(0.128)	(0.273)	(0.144)
3 - Skilled	0.354	0.403	0.317	Distributive	0.319	0.126	0.361
	(0.089)	(0.193)	(0.101)		(0.148)	(0.308)	(0.168)
4 - Partly Skilled	0.136	0.311	0.054	Mining	0.065	0.167	0.083
	(0.109)	(0.227)	(0.126)		(0.122)	(0.309)	(0.136)
Student	0.515	0.785	0.384	Textiles	-0.114	0.033	-0.200
	(0.074)	(0.162)	(0.086)		(0.102)	(0.211)	(0.119)
Father's class	(***)	(***)	(***)	Iron & Steel	-0.022	-0.049	-0.016
1 - Professional	1.427	0.930	1.630		(0.144)	(0.289)	(0.166)
	(0.159)	(0.283)	(0.203)	Other manuf.	0.204	0.271	0.192
2 - Intermediate	0.463	0.444	0.366		(0.089)	(0.189)	(0.101)
	(0.101)	(0.198)	(0.121)	*Region*	(*)	(***)	
3 - Skilled	0.342	0.213	0.348	Lancashire	0.041	-0.015	-0.037
	(0.064)	(0.142)	(0.074)		(0.089)	(0.155)	(0.120)
4 - Partly Skilled	-0.134	-0.238	-0.117	London Environs	-0.075	-0.353	0.001
	(0.065)	(0.148)	(0.072)		(0.053)	(0.109)	(0.061)
Age discrep. (years)	(***)		(***)	Wales	0.166	0.113	0.219
0	0.201	0.133	0.224		(0.116)	(0.313)	(0.126)
	(0.046)	(0.095)	(0.053)	Yorkshire	0.126	0.282	0.045
1	-0.008	-0.028	0.004		(0.075)	(0.154)	(0.086)
	(0.050)	(0.103)	(0.057)	*Moved to city*	-0.005		
Age	0.052	0.167	0.009		(0.043)		
	(0.027)	(0.059)	(0.031)	ρ_1		-0.053	
Age² / 100	-0.152	-0.505	-0.023			(0.100)	
	(0.072)	(0.162)	(0.081)	ρ_2		-0.324	
Eldest	0.030	0.111	0.002			(0.088)	
	(0.039)	(0.081)	(0.045)	k_2	0.959	0.948	
Inheritance	0.335	-0.102	0.552		(0.028)	(0.031)	
	(0.110)	(0.213)	(0.130)	k_3	2.429	2.408	
EldestInheritance	-0.039	-0.252	-0.012		(0.038)	(0.050)	
	(0.108)	(0.237)	(0.121)	k_4	3.665	3.642	
Age heaping	-0.003	-0.005	-0.003		(0.060)	(0.076)	
	(0.003)	(0.007)	(0.004)				
Living in town	0.051	-0.084	0.050	Pseudo R^2	0.092	0.090	
	(0.038)	(0.083)	(0.045)	Log likelihood	-4572.715	-6457.798	
				LR χ^2	925.530	1284.230	
				Prob > χ^2	0.000	0.000	

* = Jointly significant at the 10 percent level ** = Jointly significant at the 5 percent level *** = Jointly significant at the 1 percent level

Notes: N = 3,774. The two columns of equation (1) represent all 3,774, (2) represent the 896 who moved to a city, and (3) the 2,878 who remained in a rural place. Standard errors are in parentheses. All explanatory variables use 1851 values. Pseudo $R^2 \equiv 1 - L_1/L_0$, where L_1 is the log likelihood of the model and L_0 is the log likelihood of the "constant only" model. The LR χ^2 statistic represents a test of all parameters of the model being equal to 0. Omitted dummies are *Class*, *Father's class*: 5, *Age discrepancy*: 2-5, *Industry*: Others, *Region*: Others. k_1 is normalized to 0.

TABLE 4

MAXIMUM LIKELIHOOD RESULTS: DETERMINANTS OF JOB CLASS ATTAINMENT

Marginal effects of explanatory variables

	d[Pr(y=1)] / dx		d[Pr(y=2)] / dx		d[Pr(y=3)] / dx		d[Pr(y=4)] / dx		d[Pr(y=5)] / dx	
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Job class										
1,2 - Prof., Intermed.	0.3639	0.0921	0.3075	0.2739	-0.3423	-0.0452	-0.2437	-0.2151	-0.0855	-0.1057
3 - Skilled	0.0161	0.0081	0.0888	0.0596	0.0335	0.0515	-0.0847	-0.0636	-0.0538	-0.0555
4 - Partly Skilled	0.0120	0.0012	0.0682	0.0097	0.0273	0.0098	-0.0656	-0.0106	-0.0419	-0.0100
Student	0.0402	0.0103	0.1797	0.0734	0.0334	0.0591	-0.1594	-0.0774	-0.0939	-0.0654
Father's class										
1 - Professional	0.0811	0.2007	0.2297	0.3490	-0.0668	-0.1784	-0.1720	-0.2581	-0.0720	-0.1131
2 - Intermediate	0.0206	0.0108	0.1022	0.0726	0.0237	0.0510	-0.0930	-0.0751	-0.0535	-0.0593
3 - Skilled	0.0074	0.0084	0.0453	0.0640	0.0228	0.0591	-0.0450	-0.0690	-0.0306	-0.0626
4 - Partly Skilled	-0.0072	-0.0024	-0.0480	-0.0202	-0.0316	-0.0224	0.0494	0.0225	0.0374	0.0225
Age discrep. (years)										
0	0.0044	0.0049	0.0277	0.0396	0.0155	0.0413	-0.0279	-0.0436	-0.0197	-0.0421
1	-0.0009	0.0001	-0.0059	0.0007	-0.0034	0.0007	0.0060	-0.0008	0.0042	-0.0007
Age	0.0055	0.0002	0.0349	0.0016	0.0197	0.0017	-0.0352	-0.0018	-0.0248	-0.0017
Age ^{2} */ 100*	-0.0166	-0.0005	-0.1055	-0.0040	-0.0595	-0.0042	0.1065	0.0044	0.0750	0.0043
Eldest	0.0036	0.0000	0.0229	0.0004	0.0134	0.0004	-0.0233	-0.0004	-0.0166	-0.0004
Inheritance	-0.0031	0.0202	-0.0206	0.1167	-0.0136	0.0568	0.0213	-0.1139	0.0160	-0.0798
*Eldest*Inheritance*	-0.0065	-0.0002	-0.0477	-0.0021	-0.0402	-0.0022	0.0509	0.0023	0.0435	0.0022
Age heaping	-0.0002	-0.0001	-0.0010	-0.0005	-0.0007	-0.0006	0.0011	0.0006	0.0008	0.0006
Living in town	-0.0028	0.0011	-0.0176	0.0088	-0.0099	0.0093	0.0177	-0.0098	0.0125	-0.0094
Industry										
Agriculture	-0.0059	-0.0025	-0.0398	-0.0218	-0.0275	-0.0251	0.0413	0.0244	0.0318	0.0251
Building	0.0138	0.0092	0.0704	0.0607	0.0173	0.0400	-0.0646	-0.0624	-0.0369	-0.0475
Distributive	0.0047	0.0119	0.0276	0.0745	0.0117	0.0431	-0.0268	-0.0750	-0.0172	-0.0544
Mining	0.0065	0.0020	0.0368	0.0153	0.0142	0.0143	-0.0354	-0.0166	-0.0222	-0.0149
Textiles	0.0011	-0.0035	0.0069	-0.0322	0.0037	-0.0425	-0.0069	0.0365	-0.0048	0.0417
Iron & Steel	-0.0015	-0.0003	-0.0100	-0.0028	-0.0062	-0.0030	0.0102	0.0031	0.0075	0.0031
Other manuf.	0.0110	0.0049	0.0605	0.0362	0.0209	0.0310	-0.0573	-0.0388	-0.0351	-0.0333
Pr[y | X=E(X)]	0.0127	0.0078	0.1456	0.1105	0.5183	0.4905	0.2436	0.2807	0.0799	0.1104

Notes: Dependent variable *y* is job class in 1881. Region variables not shown. Effects are calculated at mean of *x* for continuous variables; for discrete variables, effect is Pr(*y*=c|*x*=1) - Pr(*y*=c|*x*=0), c=1,…,5. Baseline probabilities calculated at mean of all variables.

higher class in 1851 and being the son of a father in a higher class strongly predict ending up with a higher quality job in 1881. For example, Table 4 indicates that people who began in Class I or II and moved to the city were 31 percentage points more likely to end up in Class II than all others who moved to the city, and they were 36 percentage points more likely to end up in Class I. Given that the two baseline attainment probabilities are 15 and 1 percent, respectively, it was very difficult indeed to move up into a Class I or II job from a lower initial class.

Perhaps most interesting with regard to the class variables is to compare their effect on those who moved to urban areas versus the effect on those who remained in rural places. A chi-square test reveals no significant difference between the combined effect of both own and father's class on those who moved versus those who remained. However as Table 3 shows, for those who moved to a city, the effect of their own class in 1851 was stronger than the effect of their father's class. The reverse was true for those who remained behind, the job class of their father being a stronger influence on their 1881 class attainment than was their own class in 1851. Also, the effect of own class was stronger for those who moved than those who did not, and the effect of father's class was weaker, though only the former effect is statistically significant (the two χ^2 statistics are different from zero with 97 and 80 percent confidence, respectively). It appears that leaving the countryside and moving to the city offered migrants a better chance to escape the intergenerational career trajectory begun with their father's job. For those who moved, their own position mattered more than that of their father. It is also interesting to note that the benefit of being a student in 1851 was nearly twice as great for those who moved to a city as it was for those who remained rural. Students who moved to the city were 18 percentage points more likely to end up with a Class II job than non-student urban migrants; for those who remained in a rural place, the advantage was only 7 percentage points.

As for other variables of interest, people who reported their age with consistency between the two censuses performed better in the labor market. The effect was particularly strong, and statistically significant, for non-migrants, who were 8 percentage points more likely to end up with a Class II or III job and 4 percentage points less likely to find themselves in a Class V job than non-migrants who inconsistently reported their age by more than one year. Having a potential inheritance also was a strongly positive force, but only for non-migrants, as expected. The effect was particularly positive for the attainment of Class II jobs, where it increased the probability by 12 percentage points.

Self-Selection and the Determinants of Migration

Who were the urban migrants? Table 2 gives an overview of their characteristics. 896 of the 3,774 individuals in the estimation sample moved from a rural to an urban area between 1851 and 1881, for an overall urban migration rate of 24 percent. Out of the entire sample of 28,474 matched individuals (which is younger than the population as a whole because of the survival effect in the matching process), 18,740 (66 percent) were living in a non-urban area in 1851. Of these, 4,387 moved to a city by 1881, for a rate of 23 percent over the thirty-year period. As for the characteristics of these migrants, most strikingly, they were not those clinging to the bottom of the economic and social ladder, desperate for any sort of a change. We see relatively more people with Class III jobs in 1851 and sons of fathers with Class III jobs among the pool of urban migrants than among the non-migrants. Just the reverse is true for Classes IV and V, which are represented more heavily among the rural persisters. Sons who were described as students in 1851 were also more prevalent within the group of urban migrants than the group who remained in rural places. It was decidedly not the case that the truly destitute and poverty stricken, those occupying the society's lowest class, were pouring from the countryside into the cities. Rather, it was those who were middle-of-the-road—children who had received schooling, sons of fathers with skilled occupations, young men who themselves were employed in a skilled occupation—who left the rural areas of England and Wales and headed for the cities.

Table 5 presents results from the probit estimation of the urban migration decision, which yields a more accurate and detailed picture of the characteristics of urban migrants. The large, positive coefficient on $\left(\hat{y}_{1i}^{*} - \hat{y}_{0i}^{*}\right)$ indicates that in fact people were moving to the cities in order to improve their socioeconomic status and that the pull of this factor was strong relative to other factors. Interpreting the magnitude of the effect is not straightforward, as y^{*} , job quality, is a construct of the model, not an observed variable. Every unit increase in anticipated job quality difference between the rural and urban regimes increased the odds of an individual moving to a city by 7 percentage points. The estimated cut points from Table 5 shed a little light on this figure. The range of y^{*} between different job classes varies from 0.95 for Class IV to 1.46 for Class III. So the prospect of being one class higher in the city than in the countryside (roughly an increase in y^{*} of 1 to 1.5) increases the odds of migration by about 7-10 percentage points—a

TABLE 5

PROBIT RESULTS: DETERMINANTS OF URBAN MIGRATION

Variable (individual-specific)	Move to city coefficient (1)	Move to city d[Pr(*D*=1)]/dz (1)	Variable (location-specific)	Move to city coefficient (1)	Move to city d[Pr(*D*=1)]/dz (1)
Constant	-1.296	---	*Distance to city*	0.001	0.0003
	(0.501)			(0.004)	
$y_{1i}^{*} - y_{0i}^{*}$	0.239	0.071	*Wage gap*	0.035	0.011
	(0.127)			(0.016)	
Age discrep. (years)			*Previous migrants*	-4.43E-06	-1.32E-06
0	-0.055	-0.016		(-3.12E-06)	
	(0.065)		*Unemployment*	0.043	0.013
1	-0.044	-0.013		(0.013)	
	(0.073)		*Unemp*Distance*	-0.001	-0.0004
Age	0.020	0.006		(0.000)	
	(0.038)		*Nearby cities*	-0.006	-0.002
$Age^2 / 100$	-0.093	-0.028		(0.002)	
	(0.109)		*Age heaping*	0.005	0.002
Married	0.092	0.028		(0.005)	
	(0.163)		*Pct. in agriculture*	-0.685	-0.204
Not in town of birth	0.157	0.048		(0.370)	
	(0.055)		*Pct. in manuf.*	0.508	0.151
Living in town	0.281	0.083		(0.553)	
	(0.056)		*Region*	(*)	
Industry	(***)		Lancashire	0.108	0.033
Agriculture	-0.270	-0.076		(0.189)	
	(0.087)		London Environs	0.194	0.061
Building	0.025	0.007		(0.095)	
	(0.155)		Wales	-0.230	-0.063
Distributive	0.010	0.003		(0.184)	
	(0.209)		Yorkshire	0.156	0.049
Mining	-0.493	-0.12		(0.139)	
	(0.170)				
Textiles	0.056	0.017	Pseudo R^2	0.065	
	(0.107)		Log likelihood	-1933.311	
Iron & Steel	0.083	0.025	LR $\chi^2(94)$	270.290	
	(0.170)		Prob > χ^2	0.000	
Other Manuf.	-0.040	-0.012			
	(0.082)		Pr[*D*=1 | *Z*=E(*Z*)]	0.222	

* = Jointly significant at the 10 percent level ** = Jointly significant at the 5 percent level *** = Jointly significant at the 1 percent level

Notes: N = 3,774, of which 896 moved to a city by 1881. Standard errors are in parentheses, calculated by bootstrapping via data resampling with 250 repetitions, which is necessary since $\hat{y}_{1i}^{*} - \hat{y}_{0i}^{*}$ is estimated rather than observed. All dependent variables are 1851 values. Omitted dummies are *Age discrepancy*: 2-5, *Industry*: Others, *Region*: Others. Marginal effects evaluated at the mean of z, except for discrete variables, where the marginal effect is Pr(*D*=1|*z*=1) - Pr(*D*=1|*z*=0).

large increase considering that the baseline predicted probability of moving is 22 percent. Understanding the other determinants of migration is simpler. As expected, people who were not living in their place of birth and those living in towns were more likely to move to a city by 1881, by 5 and 8 percentage points, respectively. Taken together, this suggests at least the possibility of a sort of intergenerational step-wise migration, with fathers moving from the truly

rural areas to small towns, and their sons subsequently moving to the bigger cities. This is purely speculative but not inconsistent with the data. Age did not yield the usual negative effect, almost certainly because the sample is already restricted to the young. Nor did consistent age reporting yield the expected positive informational effect. Those working in agriculture and mining in 1851 were considerably less likely to move to a city by 1881.

The location-level variables were also important factors in the migration decision. The economic incentive to move was not entirely captured by the effect of job quality; the average wage gap was also a significant influence. For every shilling per week of expected wage difference between the city and the countryside, an individual was about 1 percentage point more likely to move. As anticipated, the effect of the unemployment rate was not clear-cut. The effect of urban unemployment was significantly positive, indicating that it may have been proxying for local unemployment. The test for this is to interact unemployment with distance to nearest city. The coefficient is negative, indicating that the farther away was the nearest city, the more the effect of high urban unemployment was indeed negative. Other factors, such as distance, nearby cities, and friends-and-family, exerted small and/or statistically insignificant effects.

We have seen that the urban migrants did not begin the time period as the dregs of the labor market. We turn now to the question of selection: whether the urban migrants and rural persisters were positively or negatively selected. Table 6 presents estimates for the three migration effect parameters defined in (10) and (11) .

TABLE 6
MEASURES OF THE MIGRATION EFFECT
(Dependent variable is latent job quality in 1881, *y**)

		estimate	standard error	90% Confidence Interval 80% Confidence Interval	
Selection of Urban Migrants	(s_1)	0.1337	0.0400	0.0676 0.0824	0.1998 0.1850
Selection of Rural Persisters	(s_0)	-0.0982	0.0308	-0.1490 -0.1377	-0.0474 -0.0587
Treatment Effect	(τ)	0.2087	0.1563	-0.0494 0.0082	0.4668 0.4092

Notes: Standard errors and confidence intervals are calculated by bootstrapping via data resampling with 250 repetitions.

The first two summarize the selection process. Urban migrants were positively selected while rural persisters were negatively selected. Both results are statistically significant at the 1 percent significance level. Not only did the migrants perform better in the urban labor market than the persisters would have, they also would have outperformed the persisters in the rural labor market

had they chosen not to migrate. In this sense, then, urban migrants were the "cream of the crop." They were those whose labor market prospects were brightest.[17]

Treatment Effect

One of the central questions of the present study concerns the effect of moving to a city on the ability of migrants to attain high quality jobs; i.e. the treatment effect of moving to a city. The estimate of the treatment effect, along with standard error and confidence intervals, are shown in Table 6. The treatment effect is positive. Though the estimate of τ is not statistically significant according to the standard two-tailed test, the null hypothesis that $\tau \leq 0$ can be rejected at the 10 percent significance level. Moving to the city allowed the average mover to obtain a better job than he would have been able to get had he remained in a rural place.

Magnitude is best interpreted by considering job class transitions rather than effects on the latent variable y^*. Also, it is informative to examine exactly for whom the treatment effects were largest. This information is presented in Table 7. Here the treatment effect of moving to a city is considered in terms of the effect of a move on intra- and intergenerational occupational mobility. The table gives the probabilities that an individual will attain either a higher or lower occupational class (respectively, "up" and "down" in the column headings), conditional either on moving to an urban place or remaining in a rural area. In the first five rows, moves are relative to the individual's own class in 1851, and in the second five to the father's class.

TABLE 7
TREATMENT EFFECT BY CLASS

	Pr(up|urban)	Pr(up|rural)	Pr(down|urban)	Pr(down|rural)	Net Gain
Class in 1851 = 1 or 2	0	0	0.1976	0.3874	0.1898
Class in 1851 = 3	0.2131	0.1641	0.2533	0.3148	0.1105
Class in 1851 = 4	0.5188	0.4282	0.1598	0.2215	0.1523
Class in 1851 = 5	0.8374	0.8260	0	0	0.0115
Class in 1851 = student	0.2510	0.1793	0.2153	0.2941	0.1505
Father's class = 1	0	0	0.8376	0.7432	-0.0944
Father's class = 2	0.0377	0.0341	0.7067	0.7224	0.0193
Father's class = 3	0.2225	0.1602	0.2432	0.3206	0.1398
Father's class = 4	0.5326	0.4219	0.1516	0.2264	0.1855
Father's class = 5	0.8940	0.8140	0	0	0.0800

Note: individuals' classed as students in 1851 were considered to have moved up if they obtained a Class I or II job by 1881, and they were considered to have moved down if they held Class IV or V jobs in 1881.

[17] In this aspect they appear to be different from rural-urban migrants in the 19th century U.S. Ferrie reports that urban migrants in the U.S. between 1850 and 1860 were negatively, rather than positively, selected. They fared worse in the cities than rural persisters would have had they chosen to move. See Ferrie, "Down on the Farm," p. 11.

Net gain of moving to a city is defined as

$$\left[\Pr(\text{up} \mid \text{urban}) - \Pr(\text{up} \mid \text{rural})\right] + \left[\Pr(\text{down} \mid \text{rural}) - \Pr(\text{down} \mid \text{urban})\right] \quad (13)$$

For example, the average mover who held a Class IV job in 1851 had a 52 percent chance of moving up to a Class I, II, or III job if they moved to a city and a 43 percent chance if they chose to remain in a rural place. That same person would have a 16 percent chance of falling to a Class V job in the city and a 22 percent chance of falling in the countryside. So moving to the city confers a 9 percentage point boost in the probability of making an upward move and offers a 6 percentage point lower likelihood of falling to a lower class, for a net gain of 15 percentage points.

The first important thing to note from Table 7 is that the gains from urban migration were realized across all socioeconomic strata. All individuals were more likely to improve their own socioeconomic status if they migrated to a city, and sons of fathers belonging to any class other than I were more likely to end up in a higher strata by 1881 than their fathers occupied in 1851 if they moved. The second feature to note from the table is those groups for whom the treatment effect of migration was the largest. As Table 2 shows, by far the largest groups are those who begin with a Class III or IV job or as a student; they make up 86 percent of the sample. The net gain of moving to a city was large for all these individuals. For the average student, Class IV worker, and Class III worker, the net gain of moving to a city was 15.05, 15.23, and 11.05 percentage points percent, respectively.[18] About 70 percent of the sample had fathers with Class III or IV jobs. These men realized net gains to urban migration, in terms of intergenerational occupational mobility, of 13.98 and 18.55 percentage points, respectively.

Consider, then, the average person who held a Class IV, partly skilled job in 1851, lived in the countryside, and subsequently decided to leave and move to a city. Such an individual was 21 percent more likely to improve the quality of his occupation than if he had remained in a rural area (52 versus 43 percent), and he was 39 percent less likely to fall into a Class V, unskilled occupation (16 versus 22 percent). The average rural son of a Class IV worker was 26 percent more likely to attain a higher status than his father if he moved to an urban place than if he remained in the countryside, and he was 49 percent less likely to find himself with a Class V

[18] As a group, individuals with Class 4 jobs in 1851 realized strong net gains from moving to urban areas even though many of them made the downward move from agricultural laborer (Class IV) to general laborer (Class V). But such individuals actually realized a 65 percent gain in annual earnings, on average (Tuttle, "Children," p. 173). If agricultural and general laborers were put into the same class, or if general laborers were ranked above agricultural, then the treatment effects would be even stronger.

occupation. These are substantial gains. For people like this, moving to the city was an important avenue of socioeconomic improvement.

Conclusions

Urban migrants were the cream of the rural labor market crop, in that their prospects in both the urban and the rural labor markets were superior to those of the rural persisters. The decision to move was, by and large, a fruitful one. On average, people from all socioeconomic strata who moved to the city were substantially more successful in improving their socioeconomic status than they would have been had they remained in rural areas, and they were more likely to experience upward intergenerational occupational mobility.

References

Anderson, M. *National Sample from the 1851 Census of Great Britain: Introductory User Guide*. University of Edinburgh, Dept. Of Economic and Social History, September 1987.

Armstrong, W. A. "The Use of Information about Occupation," in E. A. Wrigley, ed., *Nineteenth-century Society*. Cambridge: Cambridge University Press, 1972.

Baines, D. *Migration In A Mature Economy: Emigration and Internal Migration in England and Wales, 1861-1900*. Cambridge: Cambridge University Press, 1985.

_____ "Population, Migration and Regional Development, 1870-1939" in R. Floud and D. McCloskey, *The Economic History of Britain Since 1700*, vol 2. Cambridge: Cambridge University Press, 1994.

Booth, C. (ed.) *Life and Labour of the People in London*, 9 vols. London, 1892-97.

Boyer, G. R., and T. J. Hatton. "Migration and Labour Market Integration in Late Nineteenth-Century England and Wales." *Economic History Review* 50, no. 4 (1997): 697-734.

Ferrie, J. P. "How Ya Gonna Keep 'Em Down on the Farm [When They've Seen Schenectady]? Rural to Urban Migration in the U.S., 1850-70," prepared for the Cliometrics Conference, May 1999. [http://www.faculty.econ.nwu.edu/faculty/ferrie/papers/urban.pdf]

Great Britain, General Register Office. *The Registrar-General's Decennial Supplement. England and Wales.* London: H.M.S.O., 1921.

_____. *Census of England and Wales, 1921. Classification of Occupations*. London: H.M.S.O., 1924.

Hatton, T. J., and J. G. Williamson. "What Drove the Mass Migrations from Europe in the Late Nineteenth Century?" *Population and Development Review* 20 (September 1994): 533-59.

Hunt, E. H. *Regional Wage Variations in Britain, 1850-1914*. Oxford: Oxford University Press, 1973.

_____. "Industrialization and Regional Inequality: Wages in Britain, 1760-1914." *Journal of Economic History* 46, no. 4 (1986): 935-966.

Lee, C.H. *British Regional Employment Statistics, 1841-1971*. Cambridge: Cambridge University Press, 1971.

Mokyr, J. *Why Ireland Starved*. London: George Allen & Unwin, 1983.

Pooley, C. and J. Turnbull. Migration and Mobility in Britain Since the 18th Century. London: UCL Press, 1998.

Ravenstein, E. G. "The Laws of Migration." *Journal of the Royal Statistical Society* 48 (1885): 167-227.

Redford, A. *Labour Migration in England, 1800-1850*, 2nd ed. Manchester: Manchester University Press, 1964 (1926).

Royle, S. A. "Social Stratification from Early Census Returns: A New Approach." *AREA* 9 (1977): 215-219.

Southall, H. "Regional Unemployment Patterns Among Skilled Engineers in Britain, 1851-1914." *Journal of Historical Geography* 12, no. 3 (1986): 268-286.

Southall, H., and D. M. Gilbert. "A Good Time to Wed?: Marriage and Economic Distress in England and Wales, 1839 to 1914." *Economic History Review* 49 (1996): 35-57.

Tuttle, C. "The Role of Children in the Industrial Revolution." Unpublished Ph.D. diss., Northwestern University, 1985.

Williamson, J. G. *Coping with City Growth During the British Industrial Revolution*. Cambridge: Cambridge University Press, 1990.