

The Cliometric Society

October 1991 Vol. 6 No. 3

Report on the ESRC Quantitative Economic History Conference, 1991

by David Greasley (University of Edinburgh)

(Edinburgh) The first meeting of the UK group in Scotland was held September 13-14 at the University of Edinburgh and attracted 35 participants, witnessing a welcome increase in overseas scholars from Poland, France, Denmark, Germany, Ireland, and the United States. Financial support from the British Academy, Royal Economic Society, Goethe Institut, and the ESRC helped make these visits possible. For the future the group heard good news from the ESRC competition — fuller funding has been awarded for the next two years.

The first session comprised three papers on the interwar economy. James Foreman-Peck (Oxford) and Chris Hammond (Hull) opened with, "Closing the Productivity Gap: the achievements of central planning in the British electricity supply industry during the 1930s." Deploying a short run cost function approach, the authors argued increasing scale in British electricity generation raised productivity to almost American levels by the end of the 1930s. Rather than being a weakness, British institutions, in the shape of the CEB, provided a plank for closing the productivity gap by boosting generating scale. Seemingly British industrial policy was not always ineffective — the authors speculated whether Britain had a comparative advantage in collective provision. Discussion focused on the detail of the productivity estimates, and on their interpretation. Stephen Broadberry (Warwick), noting the US labour productivity advantage widened, wondered whether narrowing TFP might result from unsatisfactory capital utilization measures. Leslie Hannah (LSE) questioned the wider significance of the results. The market for electricity generation was unusual, yet since the US productivity advantage emerged before 1930, a narrowing gap in the 1930s may not be untypical.

Kent Matthews (Cardiff) followed with "Recession and Recovery in Interwar Britain: a decomposition analysis," which presented a macro-modeling approach to evaluating the sources of recession and recovery. In part the paper's inspiration was last years QEH meeting, where Dimsdale and Horsewood labelled the author's previous interwar model as "not mainstream," prompting Matthews to search for greater consensus. Arguing that interwar econometric history has yielded common ground on the supply side, Matthews adopted an "agreed" structure for supply, and a reduced form to investigate demand-side disagreement. Decomposing unemployment shifts suggested that supply shocks, espe-

(continued on page 15)

What's Inside

<i>Hughes Interview</i>	3
<i>Australian Conference</i>	7
<i>Labor Statistics Conference</i> ...	9
<i>SSHA Announcement</i>	12
<i>NBER/DAE Workshop</i>	11
<i>Letter from Europe</i>	13
<i>EAHE Conference</i>	
<i>Program</i>	14
<i>Clio 1992</i>	
<i>Call for Papers</i>	32

Editor's notes

As many of you know, I have returned from Australia more or less in one piece. I spent an enjoyable 10 months there and am pleased to report that Cliometrics is alive and well in the Land of Oz. You can see this by reading Paul Johnson's write-up of a conference at Australian National University last July. The society survived well without me. The newsletter was edited superbly by John Lyons with the assistance of Lou Cain. We have come to rely on Lou and the trustees of the society unanimously agreed that Lou should be a permanent associate editor. Looking to the future, we plan to continue to sponsor sessions at the Allied Social Science Association Meetings and our annual Cliometrics Conference in May.

We are asking to become an associate member of the International Economic History Association and we will propose sponsoring some sessions at the meetings in Milan in 1994. Further, we welcome the formation of the European Association of Historical Economics at a Conference this past July. A letter from the President of EAHE, Gianni Toniolo, and the program of their conference are printed below. We are hoping to arrange for joint memberships in the Cliometric Society and the EAHE, and we will let you know of progress in a future *Newsletter*.

In December, all Society members will receive their annual renewal letters. We hope that you will do two things for us. One, please return them promptly and, equally important, please ensure your telephone number, Fax number and e-mail address are all correct so we can provide up-to-date information to all members. We hope in the next year to publish another membership directory. Since this is an expensive endeavor, we would like it to have high value, and that will depend on its accuracy. [SHW]

The Vital One

The *festschrift* honoring Jon Hughes (Joel Mokyr, editor, *The Vital One: Essays in Honor of Jonathan R. T. Hughes*, Greenwich, CT: JAI Press Inc., 1991) was recently published. Readers of this *Newsletter* will be particularly interested in Sam Williamson's article chronicling "The History of Cliometrics." In his article, Sam, citing Davis, Hughes, and Reiter (*JEH* 20:4, 1960) defines Cliometrics: "The logical structure necessary to make historical reconstruction from the surviving debris of past economic life essentially involves ideas of history, economics and statistics." In this context, almost all the articles involve Cliometrics. Contributors to the volume, in addition to Sam and Joel, are Edward Berkowitz, Brian Binger, Louis Cain, Lance Davis, Robert Gallman, Jack Goldstone, Rolf Henriksson, Elizabeth Hoffman, Theresa Hutchins, Eric Jones, William Kennedy, Gary Libecap, Kim McQuaid, John Nye, Stanley Reiter, Richard Szostak, and Paul Uselding. [LC]

THE CLIOMETRIC SOCIETY

Department of Economics
Miami University
Oxford, Ohio 45056
(513)529-2850
Fax: (513)529-6992 Bitnet: Cliomets@Miamiu

ELECTED TRUSTEES

Jeremy Atack
Susan B. Carter
C. Knick Harley
Gary B. Libecap
Leandro Prados
Roger Ransom
Blyce Rotella
Jeffrey Williamson

Ex Officio Trustees

Larry Neal, Editor *Explorations in Economic History*
Samuel H. Williamson, Secretary/Treasurer *The Cliometric Society*

The Cliometric Society Newsletter is published three times per year

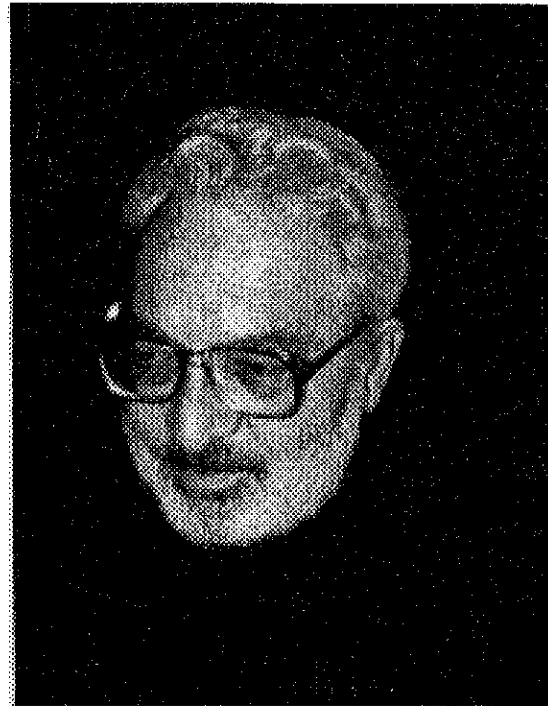
Samuel H. Williamson, Editor
John Lyons, Associate Editor
Lou Cain, Associate Editor
Margaret Voyles, Assistant Editor

Copyright©1991 by the Cliometric Society, Inc.

An Interview with Jonathan R. T. Hughes

Editor's Note: Jonathan R.T. Hughes has been Professor of Economics at Northwestern University since 1966. To most cliometricians he needs little introduction: This interview is as interesting and enjoyable as any conversation that one might have with him. It's always the same. He is a combination of story teller, scholar and renaissance man who spins a tale that you just want to listen to. And you're always sad when the evening ends. I met Jon Hughes when I was a sophomore at Purdue. I was struggling as an engineering student and thought I'd like to try economics. My concerned father said, "You ought to go talk to this professor in economics down there named Jon Hughes. See what he thinks." So I stopped by his office and he spent a while talking to me and then said, "Yes, you should be an econ major. In fact, you should go to graduate school." That's the kind of person Jon was — always encouraging, always seeing more in you than, perhaps, you did yourself. I think everyone who's known him and worked with him sees that — appreciates the way he encourages you. Later, when I became a graduate student, I had a great time at Purdue in the years that Jon so fondly remembers. I realize now that one of the reasons was that the faculty treated us students as equals — we were all scholars in the same boat. Jon was the chairman of the comprehensive committee the year I took my exams. Afterwards, he told me what he told the committee. "They all pass. You guys just figure out why." I feel that's Jon's attitude: everybody has something to contribute, so just help them along. [SHW]

The interview was conducted in the winter and early spring of 1991 by Charles Calomiris (Wharton), who was Jon Hughes' colleague at Northwestern from 1985 to 1991. Concerning his relationship with Jon, Charlie writes, "Before I met Jon Hughes, Max Hartwell spoke of him in glowing terms as friend and colleague, but warned me that I should expect to find a remarkable difference between Hughes as writer and Hughes in person. Having previously read much



of the written work, which is characterized by great erudition and close argument, I was delighted to find that his conversational style is discursive, expansive — as big as the West. In the interview that follows, I have tried to preserve in the printed word as much as I could of Jon Hughes talking, and the text, with only small variations, is a faithful transcription of how he responded to my questions and prompting. It's important to do this, I think, because to know Jon through his writings alone is to know only half of him, and at least a morsel of the other half is presented here. In the past half-dozen years I have had many such conversations with him, and more than once he was moved from talk to characteristically generous action. A few years ago he simply gave me his notes and materials from his D.Phil. research. I was in fact receiving all the secondary and primary materials on the period to date, neatly sorted and evaluated, and ready to put in a filing cabinet without even going to the archives. I'm still working through it all, and it is a mark of Jon's thoroughness as a scholar that I figure it will take me another ten years to finish with it."

Was the Cliometrics revolution a bona fide revolution or just a word applied to a slight change in technique? People like Wesley Mitchell and Simon Kuznets predate it. If it was a revolution, who reacted negatively to it and why?

We didn't think the work we did at Purdue was a revolution at first. We thought only of doing something new. We had this data-processing machinery. Purdue being the way it was, as Lance used to point out, we had computers but no library books. The steamship data were ordered from the New York Public Library. At Purdue, we had no Parliamentary papers or anything like that. If we wanted to do any kind of conventional research in those days, we had to drive to Urbana. Lance used to do that. I was told stories of him parking at a meter in front the U.I. library and putting a nickel in every hour. I think there were 250,000 books at the Purdue library when we arrived. The library expanded when the Krannert money came in and Rosenberg showed up. It was then developed into a very fancy research library in certain areas. Given the initial scarce library resources, we had to innovate, to do something with our time. We could process masses of data; the question was, what kind of data? Our idea was that we could do historical data, time series primarily, which involved calculations that were beyond the reach of manual techniques. The steamship paper had about a million computations in it. You couldn't do it on a desk calculator. As I pointed out in my "Fact and Theory" paper years ago, other people in the past had become involved in exactly the same kinds of problems. I mentioned a number of older economists going back a hundred years or so. I gave the example of William Newmarch and his bill-of-exchange survey in 1856. So we didn't think our work was revolutionary except we were using electronic machinery. It was make-work—something to do—a way of doing things that we could do there at Purdue. We only became notorious, and then "revolutionary," when Carter Goodrich asked us to talk about what we were doing at the Philadelphia meeting of the E.H.A. in 1960. That was when Stanley Reiter coined the word "Cliometrics," to put a name on our work—heavy-duty quantitative work and

data processing that hadn't been done before in economic history. I had been an accountant, and my D.Phil. thesis was filled with data that I had compiled. So I had a lively interest in the new computing machinery; I knew what drudgery we were leaving behind.

It seems to me that bringing computers to economic history wasn't the same kind of transformation of the discipline as bringing computers to I.O. or to labor economics.

The problem was to have the confidence to jump across the unknown. You had to know enough collateral historical information to give you the confidence to interpret and generalize. Old J. R. Hicks told me that he wanted to be an economic historian, but he feared the amount of unknown you had to deal with, and he preferred to be more conservative—to define his own unknowns. The problem in the early Cliometrics papers was how to go from an intuitive understanding based on theory and history to those outcomes on the printouts. I think we did provide examples, probably very imperfect ones, but examples of how you could go from what was a conventional way of thinking and take economic history into a new area. That wasn't exactly revolutionary.

Were you confident there was already an audience for that paper?

You mean "The First 1,945 British Steamships?" No, we weren't.

Who was the implied audience?

We didn't have an audience beyond Purdue. When Stan and I finished with the steamship paper, I couldn't imagine who would want to read something like that, and Stan said he wanted to try the *Journal of the American Statistical Association* because there were some things we had done in the paper which would be novel to statisticians. And that's why we went to that journal. We didn't ever send the paper around to be criticized by economic historians. What would they say? What could they have to say? They

could never have seen anything like it before. I hadn't either. We didn't know how to do any parts of it at first.

I can remember when Lance and I got the output back for the exchange rate paper. We had no idea at all what it would look like. But we got the output back and stretched it out on tables and stared at it. We realized that we had set up a machine, a bunch of equations through which these values had been fed, and they had produced this outcome. What did it mean? At that point the economic historian had to come into play because there was no way of understanding those numbers at all without the broader conception — what the economy looked like in the period that had produced those numbers. The first bunch of numbers, as is well known now, didn't make any sense. We printed them but did not comment on them. We had no explanation. The bad numbers turned out to have been the result of a setting-up error.

Do you think the newer generation of economic historians — because they are more skilled in quantitative methods and econometrics, fields that have become so much more complex — are missing some of those other skills?

I am optimistic about the profession's future.

Do you think economic history is becoming dull, less daring, less original?

I think it's too early to reach conclusions like that. You only have so much time in life to learn things. You guys can't imagine what it was like when we were graduate students, and there was no econometrics. You didn't even study mathematical statistics in most places. We had time to read history books. We had time to read the history of economic thought. We had time to think about big ideas. If you looked at Assistant Professor Morrie Morris and Instructor Dougie North, the gurus of economic history at the University of Washington in those days, that was Big Idea City. There wasn't anything else to think about. Once you get into our contemporary world, where

the graduate students are thoroughly trained in modern quantitative methods, they have the hardware readily available. You have data sets of all sorts. It's only natural that people will follow the line of highest payout, and so at first a majority of your work would look sort of picayune. Unfortunately, from the reader's point of view, many of the papers are about sets of equations and data sets. So some of the older guys of my generation are still kingpins in the world of big ideas.

I wouldn't expect this to last forever. I was terribly intrigued by the recent evidence that this country had somehow or other been malnourished between 1820 and 1860. That didn't come from anybody's big idea. I don't even think the Marxists ever claimed that. Did any contemporary ever say that in the 1880s? I don't think so. Mrs. Trollope complained Americans drank too much whiskey, ate too much hot bread, and chewed too much tobacco. The evidence for malnutrition appears first in the work of Komlos and Fogel. Now there's something there to explain, and I'm sure explanations will be forthcoming. The work seems to me to be a good example of the validity of Cliometric work; you could produce this kind of serendipity. I mean you have a big result, after all, that covered large portions of populations which nobody had ever asked about before.

If you had to guess where the \$1,000 bills are buried, what sorts of ideas, what sorts of subfields of economic history, what sorts of questions, do you think are likely to produce the next revolution?

I predicted some 15 years ago, in print, that the Great Depression of the 1930s would become the next gold mine. Apart from that kind of thing, I think modern history will force economic historians to ask questions which are not data-based in nature, but are certainly to be explained by economic theory. The largest example is the difference between the notion of freedom and fulfillment for individuals in the West compared to other parts of the world. We have just gone through a huge laboratory experiment in the application of different ideologies to economic en-

deavor. There has been an enormous apparent victory for the world of Western values, those bourgeois values which had been so denigrated by Marxist theorists for so long. I should think that a great area for big payoffs. The history of philosophy will become important again.

I was listening to two of our graduate students (one was from China and one was from Japan) talking about what was unique about the United States. They both agreed, talking only to each other with me sort of listening in, that what seemed at first odd about the United States, and then so wonderful, was freedom. Freedom from all kinds of pressures — from the family, from the society around them, not to mention from the government or from the university. Of course, this goes back to basic stuff. This has always been a big issue in the philosophy and history of political thought: why it was that Western countries had this notion that their ideal social system was one in which the maximum number of people could do as they pleased? It is not an obvious way for anybody to think of a social system. Most social systems, now and in history, are not based on such ideas. The biggest thing that's happened in the 20th century is, as the man said, we won. The question now is why? I would foresee that at first people will say it's because we have the highest per capita incomes. But that's an obvious answer, puts the cart before the horse, and wouldn't even begin to go to the root of it.

Do you think economic historians will or should focus more on the big questions of economic development?

Yes, over huge blocks of time. You know, it's perfectly clear by now there really was something acquired, say from the Greeks and the Hebrews, that made a big difference between our society and others. And the other thing, of course, is the possibility of changing the economics metaphor from physics to biology, as Mokyr has already shown. To consider human societies in a very different way and to realize that they're not necessarily going to converge at all, that they're not heading necessarily for any kind of a common end in time. Some societies will make it

farther in time, some have evolved as far as they're going to, and others contain within them possibilities for a great deal of evolution toward a longer and unique social existence.

Does history have lessons for the future?

For a long time the notion of mainstream economic development was that everybody should try to be like us. Progress was defined as Western economic structures and technology, allied with liberal democracy. That was the 19th century ideal Karl Polanyi described as utopian. One of the first papers I ever published, in the *AER*, back in 1958 (something like that), was a paper in which I argued that it was not possible, even with the kind of information available then. If you just used your head you could see it was not really possible for other countries to be just like us, because we were changing every day ourselves. We weren't a finished product. It was like trying to shoot at a moving target from a long distance: you try to be just like us, and maybe you'll get blue jeans, Coca-Cola, and nothing else. One of the points Joel made in his book is very impressive: as far as he can tell, if you want technological development, human improvement based on technological development, to be the outcome, then it's diversity you want in social organization and not uniformity.

Would you say the relationship between economic historians and other economists has changed over the last 30 years or so? And how would you describe the change?

I don't think it's changed much. The issue has always been a very simple thing: people who do economic history have an investment in the study of the past, which then becomes part of their capital, and they insist upon using it. If you don't have that investment, you don't think history is important, and obviously you're not going to use it. You don't use what you don't have. I remember, as a graduate student, some took economic history courses, and some did not. At that point a difference emerged, right there, among the graduate students in how they thought

(continued on page 18)

Conference Report: The Role of Long-run Analysis in Economics

Australian National University, July 1991

by Paul Johnson (ANU and London School of Economics)

Does history matter for economists? Does history matter for policy-makers? And should it matter? These are the questions that motivated a conference on long-run analysis organised by Graeme Snooks at the Institute of Advanced Studies of the ANU. If history in general, and economic history in particular, is to sustain the claim of 'relevance' so frequently asserted in functionalist justifications of academic endeavour, then it must show that economists can do better economics and policy-makers can make better policy if they have both a knowledge and an understanding of the past, and of the way in which past experience delimits the scope of future developments.

The conference opened with papers from Paul David (Stanford) asking 'So, how does it matter that "history matters"?' and from Graeme Snooks (IAS, ANU) on 'The lost dimension: economics without time'. David argued that the future of economics as an intellectually exciting and productive discipline lies in its becoming a properly historical social science in which the path-dependence of economic outcomes is fully recognised. The dominant ahistorical formulation of much modern economics derives from the scientific paradigms of nineteenth century economic thought which abstracted economic processes from historical events. By ignoring events in the history of economic systems, and the causal chains that events imply, economists have been able to ignore the path-dependence of economic systems and the disequilibrium conditions that must underlie any predictively useful theories of economic equilibrium. David suggested that there is already a recognition by some economists that economic theory needs to accommodate the path-dependence of economic phenomena, the importance of the timing of economic actions and the possibility of multiple equilibria in economic systems. But the

ahistorical scientific paradigms still dominate, and economic historians have an important role to play in producing detailed, persuasive examples of the way in which path-dependence has affected the evolution of economic institutions, organizations and technologies.

Snooks developed a parallel argument that the evolution of economics as a deductively-based discipline since the late nineteenth century has led to the downgrading of applied economics in general and of economic history in particular. While the development of the deductive method was essential for the technical advancement of economics, the loss of historical skills has produced a widening gulf between the technical interests of the economics profession and economic reality. In its drift from realism to abstraction, economic theory and method has become focused on the mathematically tractable concept of short-run static equilibrium. Snooks argued that for economics to regain a toe-hold on reality, and to counter the rising tide of criticism from environmentalists who claim that economics has nothing to contribute to the solution of the real problems of resource depletion, economists must begin to take time into account. If economics is to provide more realistic policy advice it must reconcile deductive reasoning with the historical specificities of long-run change.

These arguments for the re-integration of history and economics in academic studies were challenged by Bob Gregory (IAS, ANU), who suggested that economics should be viewed as a product that has become more specialised over time as the market for economics has grown. He argued that applied work is now done mainly in business and the public sector, and that universities have become specialised in the production of economic theory and technique — in

which they have a comparative advantage. That is, further segmentation and specialization seems likely for academic economics, rather than, as David advocates, transformation to a broader and truly historical social science. David countered by suggesting that specialization in economics was driven not by increasing returns but by reputation recognition which requires clear, compact and self-contained reference groups. Snooks noted that as long as reference criteria for academic economists are produced from within the group, there will be no incentive for economists to take notice of history, but external criticism of economic abstraction from environmentalists and natural scientists will demand more realistic responses.

The second session in the conference continued the critique of economic abstraction, but from the perspective of policy-makers. Gary Hawke (Institute of Policy Studies, Wellington) spoke about 'Getting your hands dirty: economic history and policy advice' and John Edwards (IAS, ANU and former economic adviser to the Treasurer) considered 'Success or failure: Australian economic policy in the 1980s'. Hawke concluded that economic understanding was a necessary but not a sufficient condition for the creation of worthwhile policy advice; also needed is an appreciation of specific institutional and historical context. Edwards was less sanguine about the role of formal economic analysis — it has a role to play in the formulation of major economic reforms, but day-to-day economic management takes little account of theoretical complexities, and is often unrelated or even antithetical to longer-run economic goals. Michael Keating (Department of Finance) took time out from the preparation of the 1991 Budget to comment on the role of economics in public policy. He noted that most public sector policy advisers are concerned with micro-economic issues of resource allocation, and that for them simple economic concepts such as opportunity cost are very powerful tools. On the other hand, the individualistic basis of economic analysis sets clear limits to its utility in considering public sector provision. Furthermore, policy advice has to recognise the way in which the range of

possible policy changes is conditioned by existing institutional structures and by expectations based on historical experience.

In the following sessions of the conference eight empirical papers were presented, considering how economic history can interact with economics and with policy formulation. Tim Hatton (Essex) presented a paper co-authored with Jeff Williamson on 'Labour market integration and the rural-urban wage gap.' Drawing on evidence for a range of industrialised countries between 1850 and 1950, Hatton and Williamson illustrated that nominal wage gaps were large and pervasive. However, when adjustments were made for the lower rural cost of living and for in-kind payments in agriculture, these gaps were substantially reduced, though the urban wage remained higher by 5 to 50 per cent. Migration flows were shown to respond to these wage gaps as well as to urban unemployment, though the response was generally not great enough to eliminate the gap. Consequently, wage gaps were not so much a condition of economic development as a symptom of unbalanced growth favouring the urban sector. They presented a model which they had estimated for several countries to account for variations over time in the wage gap. In his comments Bob Jackson (University of Queensland) noted that migration was a complex process varying across time and place and that the detailed factors affecting the propensity to migrate should be studied. David thought that the migration equation estimated for the U.S. 1920-41 might be subject to a structural break between the 1920s and the 1930s

The consequences of urban migration were considered by Lionel Frost (La Trobe) in 'Cities and Economic Development: some lessons from the economic history of the Pacific rim'. He suggested that the environmentalist arguments that many third world cities have grown beyond their ability to generate wealth, and so should be deliberately constrained in further growth by public policies, and also that the Kelley-Williamson model stressing the importance of endogenous limits to urban growth, need

(continued on page 27)

Report on Historical Labor Statistics Conference

June 21-23, 1991

by Terry Von Ende (Texas Tech University)

(Lawrence, KS) The conference opened with a paper by Martha Olney (Massachusetts, Amherst) analyzing the ownership patterns of sewing machines and musical instruments among Michigan workers in the late 19th century. There was considerable group discussion concerning the motivation behind the collection of these data—that is, what information the bureaus were trying to ascertain when they asked about these particular durable goods and the implications for what the data can tell us today. This line of questioning naturally led to the issue of why consumers bought these two types of goods and whether the variations in ownership patterns across sex, nativity, family characteristics, etc., were reasonable. Discussants felt that if sewing machine ownership was to be regarded as an investment, particularly as a hedge against unemployment or underemployment, there should be independent evidence to make the case stronger. Not everyone agreed that membership in a benefit society would be a complement. Some participants asked whether piano ownership represented the acquisition of a given level of “culture” or were purchased only for entertainment purposes? Did these motivations vary by location, say urban vs. rural? The discussion concluded with a suggestion that Olney examine the correlation between the ownership patterns of these two goods with that of other durables for later periods. A high correlation would make them indicative of ownership patterns for other durables in this period.

Elyce Rotella (Indiana) and George Alter (not present) analyzed the determinants of home ownership and mortgage holding during the late nineteenth century. The questions and discussion focused on two major points: (1) technical issues regarding the application of the life table technique to the problem of determining the “effective length” of mortgages, and (2) the hypothesis that the presence of children

greatly enhanced the willingness and ability of workers’ families to purchase homes, due to the positive effect of their earnings on future family income and, in particular, to incur debt to do so. Several participants felt that the data did not lend strong support to this latter conjecture, especially since the income variable was surprisingly weak and since another paper (Haines and Goodman) found that there was a negative correlation between home ownership and the number of wage earners in the family. There were suggestions of some alternative explanations for the positive correlation between home ownership and the presence of children. Most of these had the common theme that children alter household tastes and underscore a desire for housing that is family specific. Two discussants suggested a comparison of these data with those for contemporary Europe where rental housing contracts were different. Concerns regarding the life table technique included the problem of interpreting the possibility of dying twice (i.e. buy, sell, buy again) and the bias created by using age 65 to determine the proportion of homeowners since some parents at this age may have already liquidated this asset and moved in with children.

Day one concluded with the presentation of a paper by Michael Haines (Colgate) and Allen Goodman (not present) which estimated housing demand using data on income, personal and family characteristics of wage earners, housing type and expenditures available, using data from three of the state labor reports. The discussion began with one participant observing that boarders and renters should be treated separately since they typically represent different stages in the life cycle. Unfortunately, the data did not always permit this distinction. Several participants underscored the importance of ethnicity in this context, in part because many lending institutions

had an ethnic orientation. Many of the comments drew parallels between the value-rent ratio and the interest rate, while others asked whether capital markets should be introduced into the model since, in modern analysis, many changes in the housing market are explained by changes in the interest rate. It was also suggested that Haines consider the different implications of whether rooms were rented out as a result of a housing shortage or as a means of smoothing income. Other variables which generated substantial interest were those which probably tended to decrease residential mobility – e.g. family size, number of years with present employer, and family saving and dissaving patterns.

Bob Margo (Vanderbilt) kicked off Saturday's session with the presentation of a paper written with Claudia Goldin (involuntarily "un-present") which explores various aspects of unemployment – both voluntary and involuntary – during the late nineteenth century. Many of the initial questions asked for ways to reconcile or explain the differences between data in the census and in the state labor reports. Were there regional (agricultural vs. industrial) variations? Why was there large variation even for similar groups of workers (laborers)? This led to a discussion of whether a seasonal pattern of lay-offs could be discerned by looking at specific occupations, locations or industries, and, if so, whether the authors might be able to find evidence that employees made adjustments to these systematic shutdowns. Did a shutdown imply that all workers were sent home? Margo was also asked whether it mattered if spells of "no work" were experienced in one lump or two. This led to an elaboration on the two "types" of workers captured in these data: those that typically experienced two spells of unemployment and those that did not experience any. Others asked why the wage was not related to unemployment, why women and self-employed workers were excluded, and whether search unemployment would have been treated as voluntary or involuntary. The issue of the authors' treatment of "blanks" in the data was also addressed, as well as ways to make this methodology consistent throughout the conference papers.

The group next considered Josh Rosenbloom's (Kansas) paper which produced estimates of labor supply parameters for late nineteenth century workers, using data from Kansas. He stressed from the outset that the primary purpose was to determine how the irregularities in employment experienced by many workers affected their labor supply. Since 45 percent of the surveyed workers were union members, one participant questioned the representativeness of the sample, but another commented that the numbers were plausible since these were the peak years of the Knights of Labor and many of the workers were railroad workers. Much concern was voiced about the apparent understatement of the income effect in the estimated model and the fundamental problem of using a static model to try to capture the effects of adaptive expectations. The results also raised more general theoretical issues for consideration: (i) if the supply of labor is so inelastic with respect to wages, how can the decline in the hours worked from this period to the present be explained, and (ii) if seasonal unemployment produced changes in intensity, how did changes in the length of the working day occur? Finally, the group considered whether the categories for occupational groupings should be standardized, but decided that the most effective means varied according to the problem being analyzed.

The final paper of the morning session was presented by Chris Hanes (Pennsylvania). This paper takes issue with the typical human-capital explanations of immigrants' earnings growth, finding that, during this period, both immigrants and interstate migrants experienced slower earnings growth than in-state natives. Discussants still were not convinced that the age and age-squared variables were sufficient, asking what if age were just a proxy for a lot of omitted variables? Another participant asked what human capital really meant in this context since age and age-squared were apparently much more important than now. Yet another suggested that at the very least, Hanes should control for occupation since age might proxy human capital fairly well for native-born Americans but not for immigrants. On this point,

(continued on page 29)

Report on NBER/DAE Summer Institute

by Robert A. Margo (Vanderbilt University)

(Cambridge, MA) The Development of the American Economy Program (DAE) of the National Bureau of Economic Research (NBER) held its annual Summer Institute July 22-25. Thirty one people participated in the program, consisting of 15 papers, which was organized by Michael Haines (Colgate) and Robert Margo, with the assistance of Gary Libecap (Arizona) and Eugene White (Rutgers).

Richard Sutch (UC, Berkeley) opened the institute with a report on the NBER Historical Labor Statistics Conference, summarized elsewhere in the Newsletter. Sutch described the Historical Labor Statistics project which, when completed, will have collected information from various late nineteenth century reports of state bureaus of labor statistics on approximately 125,000 individuals. Discussion of the project focussed on issues of data reliability, characteristics of the sampled individuals, and why the reports were prepared in the first place. Following the summary, Sutch presented a paper, co-authored with Susan Carter, using state BLS data on wages and employment patterns in manufacturing firms during the depression of the 1890s. Carter and Sutch find evidence of countercyclical labor productivity in the 1890s, opposite to the modern pattern.

The next two papers, presented by Robert Fogel (Chicago) and Dora Costa (Chicago), concerned a new DAE project, "Early Indicators of Later Work Levels, Disease, and Death." The database for the project consists of 10 interlinked sources giving life-cycle information (15,000 variables) on approximately 40,000 Union Army veterans. Discussion of Fogel's paper focussed on problems of record linkage, sample selection bias, and the reasons why economic historians should be interested in the effects of wartime stress (for example, wounds received during battle) on economic behavior later in life. Costa presented a preliminary analysis of mortality using the life-cycle data, finding significant

effects of height (measured at enlistment) on mortality risk later in life.

Tuesday, July 23, was "macrohistory" day, now a regular feature of the Summer Institute. Hugh Rockoff (Rutgers) opened with a paper on Gresham's Law, co-authored with Robert Greenfield. Recent work by Rolnick and Weber suggests that Gresham's Law did not apply in the nineteenth century United States. Rockoff and Greenfield reconsider the episodes analyzed by Rolnick and Weber, and conclude that a modified, "93 percent" version of Gresham's Law still belongs in the monetary economist's tool kit. Discussion heated up as Michael Bordo (Rutgers) and Charles Evans (South Carolina/FRB Chicago) next presented their paper, which developed an equilibrium real business cycle model of the Great Depression of the 1930s. Although participants thought that the Bordo-Evans paper was a useful exercise, there was skepticism about their conclusion that a large, exogenous increase in the cost of transacting was a useful way to interpret economic downturn in the early 1930s. Eugene White and Peter Rappoport (Rutgers) presented a paper analyzing individual stock prices in the late 1920s and early 1930s. White and Rappoport found evidence of "herd behavior" (movements in stock prices that could not be explained by economic fundamentals) after, but not before, the downturn in 1929. Macrohistory day ended with Richard Sylla's (NYU) paper, which provided new annual series on stock and bond returns from the early 19th century to the present. Based on trends in the series, Sylla dated the onset of modern economic growth back to 1815.

Wednesday, July 24, was devoted to political economy. John Wallis (Maryland) opened with an expansive presentation entitled "Form and Function in State and Local Government." The goal of Wallis' research is to explain why the state's aggregate share

of state and local government spending increased markedly in the twentieth century. Jeremy Atack (Illinois) followed with a paper (co-authored with Fred Bateman) on the effects of protectionist legislation on hours worked in the late 19th century, using data from the census manuscripts of 1880. Their basic finding is that the laws had no effect on hours in 1880, except in Massachusetts. Shawn Kantor (Arizona) next presented some preliminary results on his joint project with Price Fishback on the passage of workmen's compensation legislation in the early 20th century. Kantor and Fishback find that passage was more likely in states in which the risk of accidents was high, but was less likely in states with a large mining or agricultural sector, holding accident risk constant. Wednesday's session came to a close with a provocative paper by Ed Ray (Ohio State), which examined how the U.S. tariff structure has affected global pollution levels in recent decades.

The final day of the DAE summer institute focussed on labor history and related topics. Robert Margo presented his paper, co-authored with Al Finegan, on

labor force participation of black male teenagers in the South from 1900 to 1970. Margo and Finegan demonstrated that the post-1950 decline in participation, frequently attributed to mechanization of cotton agriculture and the minimum wage, was actually the continuation of a long-term downward trend driven by rising school levels. Mike Haines presented a paper on homeownership using late 19th century state bureau of labor statistics reports, which he had given at the Kansas conference. The institute came to close with Richard Steckel's (Ohio State) analysis of attempts to assess underenumeration (and other errors) in the nineteenth century federal censuses. Nearly every participant had a horror story to relate about errors discovered in the census manuscripts or published census volumes.

A full catalog of all papers presented at the various NBER summer institutes will be available from the NBER later this year. Copies of DAE summer institute papers included in the NBER working paper series may be obtained from the NBER upon publication in the series. Copies can also be obtained from each author.

ANNOUNCEMENT

The Social Science History Association will hold its annual meeting November 5-8, 1992 in Chicago, IL. Paper proposals are due by February 15, 1992.

Those interested should contact:

Margo Anderson
Department of History
University of Wisconsin — Milwaukee
Milwaukee, WI 53201
414-229-4361
margo@csd4.csd.uwm.edu

or
Andrew Abbott
Department of Sociology
University of Chicago
1126 E. 59th Street
Chicago, IL 60637
312-702-4545

Letter from Europe

July 31st, 1991

Dear Friends,

Rungsted is located approximately half-way between Copenhagen and Elsinore. It is as clean and tidy as only Scandinavian (and Swiss) small towns can be and is known to the outside world for being the home-town of Karen Blixen. People there seem to care for their own history: the house of the great novelist has been turned into a lovely museum where passages from *Out of Africa* are read to visitors and one can almost see Babette preparing her great dinner.

Rungsted has now a new claim to world-wide fame. It will be remembered (at least by a handful of cliometricians) as the birthplace of the European Association of Historical Economics. Yes, dear friends, what was conjured up last year in Leuven, with the help of some of you, is now official. The Association has been formally established by some forty-odd funding members gathered at the Rungstedgaard on July 19-21, 1991 for their first Conference.

The first President of the EAHE is Gunnar Persson of the University of Copenhagen, certainly known to you for his cliometric work on pre-industrial economies. President elect is Rainer Fremdling, also a well-known cliometrician, who will try to organize the next Conference in Groningen (Netherlands) two years from now, in the Summer of 1993. Treasurer and editor of the Newsletter is James Foreman-Peck. Those of you who like to join should send him 10 (British) pounds (five only if graduate student). Trustees are: Leonid Borodkin (Moscow), Claude Postal-Vinay (Paris), Leandro Prados (Madrid), Jaime Reis (Lisbon), and Gianni Toniolo (Venice).

Why "Historical Economics" and what will the Association do? As for the name, we had decided from the beginning against "European Cliometric Society." There is already a World Clio and we can't

stretch our Muse too much. Moreover, it seemed to some of us—including myself—that the term tends to be perceived too narrowly in Europe for the Association to achieve its intended goal: the broadest possible membership of all those who apply the tools of modern economics to the study of the past. Here in Europe we may draw on a time-honored tradition of scholars trying to make explicit use of economic theory (any economic theory) in writing history dating at least as far back as the German Historical School (but didn't most of the classical economists use their economics to explain the past as well as the present?). Such a tradition is divided into a variety of streams and paradigms having little in common except the awareness that the past, like anything else, can only be explained with the use of precise theoretical tools. This was the methodological message of such scholars as Weber, Schumpeter, Heckscher, Einaudi, Dobb, Hicks. Measurement is an important but not a necessary component of this kind of economic history. A number of friends proposed "European Association of Analytical Economic History:" a ballot was taken and "Historical Economics" prevailed. I hope that this choice will strengthen our links with the economics profession: an important goal here in Europe (where, as you know, economic historians are mostly non-economists) and one of the Association's objectives.

What are our plans for the future? First of all, we find it important to establish formal links among European "historical economists," to exchange information about research projects, seminars, conferences, scholarships and the like. In order to do so, we plan to produce a Newsletter. The second aim of the Association is the organization of one official conference every other year. We feel that we do not have enough resources to plan annual gatherings, at least for the time being. We shall, however, co-sponsor workshops on specific issues whenever possible and

PROGRAM OF THE FIRST EAHE CONFERENCE

Banking: Central, Free and Regulated

Forest Capie and Geoffrey E. Wood: "Central Bank Dependence and Performance: An Historical Perspective"

Michael Collins: "Central Banking in History: Lessons from Britain"

Jose Luis Garcia Ruiz: "Free Banking in Spain 1856-1874: A Revision"

Per H. Hansen: "Bank Regulation in Denmark up to 1930"

Jaime Reis: "Monopoly Versus Competition in the Supply of Paper Money under Convertibility: Portugal, 1860-1890"

Problems in Italian Industrialization

Carlo Bardini: "Energy Economy and Industrialization: New Perspectives upon a "Crisis" Period (Italy 1886/89-1894/95)"

Alberto E. G. Schram: "Assessing the Impact of Railways in the Veneto Region"

The Future of Economic History, Two Views:

Rolf Dumke and James Foreman-Peck

All That Growth

S. N. Broadberry and B. van Ark: "Convergence and Divergence of Comparative Productivity Levels in Manufacturing: A Long Term Study of Europe and the United States"

Niels Kærgård: "Macroeconomic Methods in Economic History — survey based on cases from Denmark"

Angus Maddison: "A Long Run Perspective on Saving"

Victorian Times

William P. Kennedy and Robert Delargy: "The Finance of Innovation in Victorian Britain: The Case of the Electrical Industry"

Giovanni Federico: "An Econometric Model of World Silk Trade 1870-1913"

Pedro Lains: "Growth, Exports, and Causality in Western Europe, 1865-1913 (with special reference to Portugal)"

Interwar Economics: Germany

Albrecht Ritschl: "New National Accounts for Interwar Germany"

Arthur van Riel and Arthur Schram: "Weimar Economic Decline, Nazi Economic Recovery and the Stabilization of Political Dictatorship"

Interwar Economics: Small Countries

J. W. Drukker: "Small Countries during the Great Depression or Another Look at Unemployment"

L. Schön: "Capital and Electricity in Swedish Structural Change up to the 1930s"

Government and Regulation

J. Confraria: "Monopoly Power and Production Efficiency: A Study of Portuguese Industrial Policy (1931-1974)"

Jean-Pierre Dormois: "Comparative Productivity and Efficiency in Government in Britain and France on the Eve of World War One"

Rainer Fremdling and Günter Knieps: "Competition, Regulation and Nationalization, The Prussian Railroad System in the 19th Century"

Agriculture: Prices and Institutions

Leonid Borodkin and Mikhail Svishchev: "Macroanalysis of Social Mobility in the Soviet Private Economy: Simulation of Peasantry Dynamics of 1930s"

Jean-Michel Chevet and Pascal Saint-Amour: "Market Integration in France's Wheat Sector in the Nineteenth Century"

Francesco L. Galassi: "Tuscans and Their Farms: The Economics of Share Tenancy in Fifteenth Century Florence"

Reinterpretations of the Industrial Revolution

Patrick O'Brien

organize some of our own if the Community, as we hope, provides some funding. Another possible area of activity — still under consideration — should be aimed at graduate students. In this part of the world, with the exception of some fortunate enclaves in Britain and Sweden, prospective economic historians are brought up by History Departments. The Association would perform a most useful task in organizing workshops, full-immersion courses, summer schools and the like, in order to train graduate students in quantitative techniques, econometrics and economic theory. We are studying both the intellectual and the resource sides of the problem.

As for the conference itself, I must say that its outcome exceeded my most optimistic expectations. The quality of the papers, though uneven, was on average most satisfactory. A few papers were really outstanding. I was particularly impressed by the "relevance" of most research areas (remember what John Hicks had to say about the need for economists to choose "relevant" topics?). And, of course, old Europe being still culturally so divided, one of the

intellectual kicks of our first gathering was the diversity of approaches, view-points, and choice of topics. We spoke simple English (the modern equivalent of the elementary Latin, so different from that of Cicero, used by scholars and churchmen in the Middle Ages) but the intellectual exchange was diversified and rich. The event was socially pleasant, too. Of course we missed Don's guitar and cans of Clio oil but the atmosphere was as friendly as the one enjoyed at the US Clio meetings. And Gunnar's organization was simply terrific.

As a member of the Cliometric Society, I can only hope for the swift development of cross-membership in the two organizations and I am looking forward to a World Clio Meeting co-organized by the two of them.

Gianni Toniolo
Dipartimento di Scienze Economiche
Ca' Foscari
Università di Venezia
30123 Venezia, Italy

ESRC Report *(continued from page 1)*

cially benefits, accounted for 35% of unemployment to 1925, whereas demand shocks were more important 1925-30. Subsequent depression and recovery had domestic and external elements. In the discussion Nicholas Dimsdale (Oxford) agreed supply was important, but suggested productivity and the terms of trade should be stressed rather than benefits and union membership. Generally however Matthews' views attracted little dissent. As chairman Broadberry noted, peace appears to have broken out among interwar UK macro-modelers.

Pascal Griset (IHMC-CNRS, Paris) completed the first session with, "Patents Used as Historical Data: the case of radio technology." Griset described French patenting legislation and the information the system generates, focusing on radio transmitter patents. Total numbers of patents and the nationalities of patentees were reported for the period 1899-1939.

The main expansion of radio patents occurred in the early 1920s with French nationals chiefly responsible. Subsequently there were increases in the proportion of patents granted to other nationals, Germans especially. The discussion chiefly focused on interpretation of the data. In response to Greasley, Griset argued military contracts helped shape the rise and decline of French radio technology. Hannah suggested the data might be useful in testing for international comparative advantage. On the reduced US patenting activity in France after 1914, Griset suggested this stemmed from co-operation between French and US firms.

The second session was devoted to demographic and urban history. Bernard Harris (Southampton) commenced with "The Demographic Impact of the First World War: an anthropometric perspective." Harris's chief concern was to dispute Jay Winter's

thesis that the years of the First World War saw sharp improvements in civilian health. By considering both infant mortality and the heights of children for British regions between 1900-30, Harris argued that continuity was the chief characteristic. The interwar years, not World War One, witnessed the biggest improvement in health standards before 1939. The issue which most puzzled the group was why the War's effects should appear neutral. Were there offsetting influences? Richard Anthony (Edinburgh) suggested there were beneficial improvements in diet as meat consumption fell. Roderick Floud (City of London Polytechnic) argued that there are more important influences on health than diet; both the home and work environment need to be considered.

Richard Tilly (Münster) continued with "Municipal Socialism and Municipal Enterprise in Germany, 1890-1912." Noting that German big city growth was the fastest in Europe 1870-1914, requiring rapid improvements in education, health services, and transport, and provoking a collective response chiefly via municipal government, Tilly sought to explain the motives and goals of municipal enterprises. Was municipal enterprise growth a response to market failure, a mechanism for limiting local taxes, or did it stem from the self-interests of local bureaucracy? Tilly pursued these issues by considering municipal pricing policies, and by pondering the connection between municipal enterprise revenue and the tax burden. Tilly argued municipal activity did not lighten the tax burden, although missing data was a problem. On pricing, Tilly focused on city gas works, estimating cost functions analogous to those Millward used for British gas utilities. It appears German works did not expand output much beyond the profit maximizing level. Discussion focused on model specification, the data, and interpretation of results. Broadberry suggested tax burden rather than municipal revenue should be the dependent variable. Peter Wardley (Bristol Polytechnic) reckoned the problem of missing data was less severe than Tilly thought. On interpretation Robert Millward (Manchester) and Foreman-Peck noted the cost function estimates might indicate output rather than

profit maximization.

Jeanette Brock (Strathclyde) concluded the Friday sessions with "The Importance of Emigration in Scottish Regional Population Movement, 1861-1911." Outmigration rates from Scotland were third highest in Europe during this period. What was unique about Scotland, according to Brock's estimates, was that emigration was predominantly from urban areas. Dudley Baines (LSE) noted his research, using similar techniques, found typically rural origins for English emigrants, and most discussion focused on explaining this disparity. James Heskin (UC, Dublin) suggested the temporary movement of the Irish to Scotland might be the source. Internal Scots migration might also cause bias, though Brock had adjusted for this. Anthony raised the issue of the emigrants' motives, and Brock argued Scots emigration was not driven by destitution but by choice. Skilled workers were pulled to the New World.

Some of the Conference's liveliest debate arose in the first Saturday session on the late Victorian and Edwardian economy. Two papers, by Katherine Watson (Oxford), "In Search of a Culprit: iron and steel in the UK capital market 1870-1914" and by Bill Kennedy (LSE), "The Finance of Innovation in Victorian Britain: the case of the electrical industry," dealt with capital market efficiency. Watson stressed the relation between firms and the capital market could take many forms, depending on the nature of the business. Volatility in the iron and steel market was reflected in equity prices, and made financing steel investment by new issues problematical. Steel companies did use the public market, but issues were erratic, and debentures less common than in the more stable brewing sector. Watson was uncertain whether the capital market impeded steel investment, but felt the influence of financiers was not paramount. Kennedy, on the basis of the electrical industry's experience, was more critical of the capital market, since establishing new firms simply by ordinary share capital was difficult. Kennedy argued that financial support was nearly as important as technological capability in the early electrical indus-

try, given the front-loaded capital needs, technological complexity, and the importance of scale. Reliance on stock exchanges issuing ordinary shares, given the problems of volatility and manipulation, meant there was less support for British than overseas electrical firms. The discussion centred on how the appropriateness of capital allocation should be assessed. Hannah argued that unwillingness to supply firms with capital does not indicate market failure. Watson agreed, noting the full circumstances surrounding investment decisions should be considered. Kennedy differed, stressing missed opportunities, and consensus was not achieved. In relation to iron investment Wardley agreed that output was volatile, but reckoned non-perishable stock would reduce the risk element. Foreman-Peck felt that differences in debt-equity ratios between industries would reflect risk.

John Dodgson (Liverpool) completed the session with "British Railway Cost Functions and Productivity Growth, 1900-12." Dodgson's main aim was to assess the benefits of railway amalgamation, using statistical cost estimation techniques to measure returns to scale and density, and to estimate productivity. Returns to scale were found constant, and those to density negative. Productivity growth, on the most optimistic estimate, was very modest. Part of the discussion focused on the negative returns to density. Baines was skeptical of this result, noting that greater speeds might have given better capital utilization. Foreman-Peck wondered why productivity growth was so poor. There were company differences. Why couldn't other companies emulate the achievements of the North Eastern Railway?

The final session was devoted to non-British topics. In the absence of Leandro Prados (Madrid) due to illness, the session was restricted to two papers. Firstly Wojciech Roszkowski (Warsaw) presented "East European Land Reform Between The World Wars." Roszkowski assessed the economic effects of land redistribution by considering the productivity of large and small farms. Generalizations for eastern Europe are difficult since in some states, e.g. Romania, plots were tiny, while in the Baltics average farm

size was comparatively large. As a rule large estates had higher yields than small farms, although in response to John Singleton (Manchester), Roszkowski noted the intensity of animal breeding was greater on smaller farms. Discussion centred on whether better technology or labour incentives were responsible for productivity differences. Roszkowski noted that land quality may have been worse on smaller farms. Millward wondered whether total factor indicators might show different results to the partial measures deployed in the paper.

Bill Hutchinson (Miami) completed the proceedings with "United States Regional Growth and European Trade, 1870-1914". Hutchinson reported estimates detailing the regional basis of US overseas trade. Considering seven US regions and six European countries sheds light on comparative advantage and the foundations of late 19th century US development. A preliminary econometric explanation of net US exports was also offered. In response to Greasley, Hutchinson noted export-output ratios were well above the 6% US average in some regions, the South especially, and overseas trade was important for regional growth. Both Foreman-Peck and Broadberry welcomed the data and the attempt to elucidate the US comparative advantage, but hoped that the effects of tariffs could be incorporated.

The next Quantitative Economic History Conference will be held at St. Antony's College, Oxford in September 1992 and will be organized by James Foreman-Peck.

Hughes Interview *(continued from page 6)*

about economic problems. Suddenly a homogeneous collection of graduate students seemed to have divided themselves up between humanists and "engineers." Suddenly there were areas in which we no longer could talk to each other very well because there was silence from those who took something else besides economic history in the program. In my case, at the University of Washington, some students opted for urban planning instead of economic history, and our paths separated forevermore. We all had exactly the same training in theory, but if you had this different experience that entered into your hard disk as part of your thinking material, then you were not the same as those who didn't know anything about it, and this made a difference. Later on, those who had stuck with theory, and only theory, developed an interest in techniques that went beyond matrix algebra, differential calculus, and mathematical statistics, and that then became a very heavy specialization on their part which was the equivalent of the specialization of economic historians on the historical side. So this made the range of information much larger between the two groups of people, but I don't think at the intersection it made any difference.

It always was very valuable to me, and at times very amusing, that Stanley Reiter and I always worked so well together on specific problems. The intersect was always the same, the tangency was always the same, and yet we were always bringing material from vast distances in each of our universes to bear on problems. In one case we came up with a paper on the law and using the law to study regulated economies. It has been argued that no one single living person would be able to read our paper except Leo Hurwicz. I always thought of that paper as a duet. Each of us brought our own skills to the music. It is like the Bach "Two Part Inventions" which sum up to three distinct pieces of music: A, B, and C, the blending of A and B.

This brings me to what I would describe as my Purdue questions. Is that part of what made the group of people at Purdue unusual in comparison

to other departments? A sort of willingness to find tangencies? If not, what did make that group so special?

Purdue was an example of the rule of a benevolent despot. The man who ran Purdue was E. T. Weiler. He was a student of Arthur Marget at Minnesota. Weiler had been in the Federal Reserve System; he'd been at the University of Illinois. He didn't like the economics profession. He thought the economics profession was a fraud and irrelevant, and so he began building Purdue which had no real tradition of its own. Weiler had remarkable taste for talent and hired the group which became so famous. He was very explicit about not wanting to see economics with a capital "E" reappear at Purdue University. For a long time macroeconomics was not even taught there because it was thought to be too prosaic. We had no one doing public policy because Weiler didn't want that around. Weiler hired people by himself for a long time; all those famous people at Purdue were hired as young assistant professors or even instructors by Weiler on the basis of his own intuition. That kind of success rate is fairly unbelievable. Weiler believed people who were trained in theory and did empirical work were going to advance the discipline. We developed a certain perverse pride doing things differently, and a number of novelties came out of there because of that. Because of that atmosphere of "anything goes," there were no holds barred. Nothing was going to be illegal.

This relates to your point about Western economic development and diversity.

Well, yes. You thought of that, I didn't, but that's what Weiler was after. He told me once, "You know, Jon, I have a recurrent nightmare, and the nightmare is that I will end up running a good Midwest economics department."

I remember when I told him about writing *The Vital Few*. He had heard about my activities. In those days the idea of writing about entrepreneurship seemed incomprehensible. I had only recently become a full professor. He invited me out to his house one

Saturday in the winter, and we had some drinks in front of the fire. He asked me what I was doing. I explained it to him, and he stared at me for a minute when I finished talking and then said, "Well, Jon, you're a full professor now and you can do anything you want." When *The Vital Few* came out, he bought 85 copies of it to give to his friends. That was the way he was, you know.

I had this bee in my bonnet. Right after we had launched Cliometrics in 1960, I was already thinking about *The Vital Few*, and I was no longer thinking about Cliometrics. We had made the first big steps in Cliometrica. They were great fun, but I hadn't left the world of finance and gone into academic life to do the same damned thing over and over the rest of my life. I had this new idea. I remember telling Doug North about it at the first Clio meeting, and him saying, "What a neat idea for a book." Doug was another person who would do almost anything to avoid having to hear the same story twice.

Is it possible to have Weiler's kind of entrepreneurship at a major university today?

You couldn't do it where there was already a bunch of full professors or where there was already a tradition. Purdue was a big university, a lot of resources, and a very small and insignificant economics department, so Weiler was able to just set the old profs aside, tell them to enjoy life. He even got them some part-time jobs outside of Purdue so they could make some extra money. He told them they didn't have to play our game, which was publish or perish, up or out.

Purdue became a kind of hothouse. It was so sensational that within 3 or 4 years after Weiler started, people from the Ford Foundation came there to investigate. They were convinced there was some secret research technique or organization that was causing this one department to make such a ruckus in the profession. In the end they never could find the reason for it. They gave us a big grant anyway. But I remember them, coming around the offices and wanting information which might explain why there

were so many novel and revolutionary contributions at Purdue. Weiler's methods worked. Some guys would score again and again like Vernon Smith, who was simply an extremely creative guy who had been turned loose and supported. Everything he touched changed the profession. Finally a whole building was tailored for his experimental economics, once he got going on that.

They built a building?

Partly, the building had rooms built into it for Vernon's experiments, built into the structure of the buildings, one-way mirrors and all kinds of stuff.

The Ph.D. program itself, when you look at the students who were turned out of the Ph.D. program, many are famous people. Why were they like that? Well, again, in part it was that they didn't have to learn anything simply because it had existed for a long time. The major theoretical stuff that existed at Purdue had just been invented, and that's what they were trained in. Purdue became a collection of extraordinary personalities and finally, for reasons which are very diverse indeed, it all fell apart in about two or three years.

Once it started...

Everybody left. I've forgotten the number. Something like 19 full professors got out of there in 2 or 3 years.

Why couldn't Weiler stop the exodus?

I don't think he wanted to. He had gotten interested in the management school, and he thought the management school had more to contribute, I think. He was happy with the management school. If you talk to people who left Purdue you'll hear all kinds of different reasons. There had been two or three of these places like Purdue in the history of economics, both here and abroad: Iowa State, the University of Birmingham in England. You need a certain combination of things, including the lack of constraining forces, like an old department with a lot of full

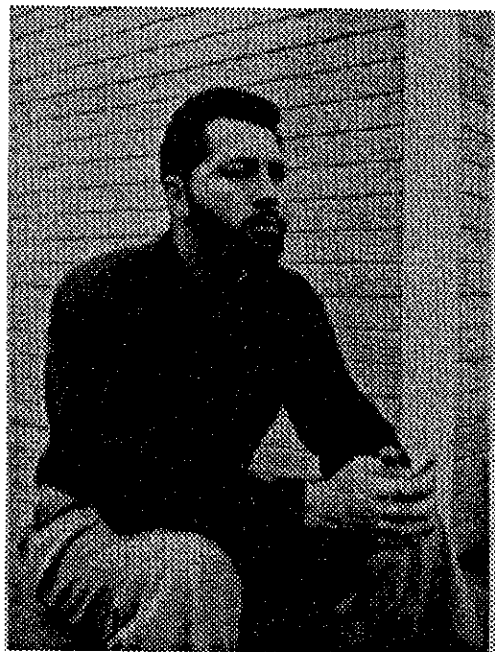
professors, tradition and stuff, that determines where anything is going to be able to go.

One thing remarkable about your work is the variety of audiences you've written for and the variety of types of books that you've written, from your Oxford product on Britain in the 1850s, to *The Vital Few*, to *The Governmental Habit*, to the work on the Japanese internment camp, to Davis and Hughes, Hughes and Rosenberg, and your work with Stan Reiter. All are very, very different. Having gotten to know you, I think there's a relationship between the variety of styles and variety of lives. You worked at the New York Fed for a while, you worked in a salmon canning factory in Alaska, you were at Oxford, you grew up in Idaho, you helped construct a Japanese internment camp, you travelled all over the place. I would guess that you must have felt like something of a chameleon in your personal life. Do you think that had an effect on your diversity of styles and subjects?

Well, I always thought of my heterogeneousness of social roots a great advantage because I was able to relate to all sorts of people. Some of the work I did that seemed to categorize me, not in one category but in several different categories, represented an overt attempt to do something different. I'll give you two examples. My Oxford thesis was an attempt to satisfy Habakkuk's desire for me to get a D.Phil., and therefore represents stone-fisted research and very careful discussion.

The Vital Few was an overt attempt on my part to correct what I thought had been a bad mistake in the study of entrepreneurship: the absence of the personalities of the entrepreneurs. Because of growing up in Idaho, I was well aware that without entrepreneurs

you couldn't have capitalism, and I knew that long before I ever read Schumpeter because I grew up in a town that began rising out of the sagebrush in 1907. So the town was only 25 years old or so when I became cognizant of what was going on around me, and I knew from looking, that businesses were entrepreneurial enterprises that hadn't had time to fossilize into corporations. Where there were people doing things, making things, and hiring others to work for them, there were entrepreneurs. So I had a vivid and lively notion of where the American economy had come from to such an extent that, when I began learning other analyses (like neoclassical economics, Marxism, and so on), to me they were all



J.R.T.H. at Nushagak, Alaska, 1952

just fruitless mental exercise. Nothing was said that would cause a man from Mars, say, to understand where this thing came from. So, when I wrote about the entrepreneurs, I wrote about it partly to make a point, and partly to have some fun with my writing. I became an academic because I wanted to write. I wrote *The Vital Few* expressly to reach a large audience.

There again, when you describe these feelings about *The Vital Few*, you're making me think about your presidential address on the Alaskan canning factory

where the point, in many ways, is the same. The point is that you have to understand the sort of individual experience of being a worker and an entrepreneur in a particular place to understand what was happening.

That's the trick, isn't it? I was just reading Peter Medawar recently, his book on philosophy of science. There's this problem about induction and deduction. If you rely on deduction alone, you have a general view of the earth and probably no information at all concerning the species on it. Even if you

become a great inductivist, on that basis alone you have no idea what the earth looks like overall, although you'd be really great on, say, red ants or termites. And the trick always has been, and it's an art, to go from specific to general knowledge. Either you have to do it, or the specific knowledge just remains in academic journals of interest to those who are interested in red ants and termites. You have to make the jump, and it's a very risky thing. There's some point at which you become convinced you may be wrong, but then you become convinced that, hey, I've got something here that's of general interest. Then you go ahead and produce it. It is a risk, and I don't think you can avoid it.

How did *The Governmental Habit* fit into your plans as an academic entrepreneur?

My book, *The Governmental Habit*, is an overtly polemical effort, but as such it is a failure. I will cause no changes in the real world with this book. I'm glad I did it, and it's been worth doing. It's very interesting to see, in small detail, where the American habits of regulation came from, but I despair of the possibility that the amount of regulation in this country will ever be reduced, whether we know anything about its origins or not. As a historian, of course, I just naturally think that our decision would be wiser if we knew why we were making it. The economy, as it exists, is an artifact. There isn't a thing that doesn't represent the past. All the rules, all the regulations, all the institutions, all the physical capital, the social structure and everything else is nothing but an artifact of the past, and to my mind it is quite pointless to think about regulation and property rights without finding out what were the property rights, where did they come from, what's the origin of our present set of property rights, and where did these methods of regulation come from. All you have to do is live in some other country and you realize that what we do is quite different from what other people do. We have our own way of regulating path dependency: it represents what we did in the past. We could easily improve our system if we would be willing to learn from anybody else, but like others, we don't. We do things mostly as we did in the past.

Does economics provide a foil for economic historians? Do economists provide encouragement because their models often are incomplete or wrong?

Well, they make very easy targets. But, after all, as an economist, in many respects my thinking is just exactly like everyone else's. For example, I always turn off the TV when an economist comes on, and there's a reason for this.

Why?

Because I know he doesn't know what he's talking about. The reason he's there is because he doesn't know what he's talking about, and therefore he can reach a broad audience. But you have to think with economic theory, you haven't got anything else to think with. I've always said theory is like a flashlight in a forest. It's useful for illuminating some points, but only if you know where to look and what you're looking for. When I see economic analysis that is based on not understanding or misunderstanding the historical background of what's being discussed, I feel very sad. There aren't many professions in which the subject's history is of no interest to the practitioners. I mean, medical people are interested in the evolution of mammalian bodies, physicists are interested in geology, astrophysicists are interested, after all, in the most historical thing there is — the origin of the universe. Most serious kinds of study, musicology, for example, are interested in the change in the nature of the thing being studied over time. I think there must not be many professions that are as unchanged in their basic theoretical thinking by changes in the facts as economics is. I don't expect the basic core of thought in economics to ever be changed by any new discoveries about the economy itself.

Is that, in a sense, something that defines economic historians? By virtue of their difference along those lines?

Well, it's been kind of pitiful. You see efforts made again and again to show that what was found in fact

was justified in theory. In those cases, one thinks, how sad; as I said long ago in the "Fact and Theory" paper, if I see a six-toed horse I should be interested in that for its own sake. I should not ignore it because theory says there are no six-toed horses. If I find one, I should report it even if there is no explanation at all. A good example of this, return again to Komlos and Fogel, is they have reported this decline in the stature of average American males when no one ever asked them to find that information and we don't know why it happened.

You have just come out with the third edition of your textbook. How did you get started on it and how did it change your views of your colleagues' work?

I had been teaching economic history for 25 years, so I had already assimilated the work. I read the journals to keep up, so I knew before I started writing it how it was going to come out. Not that it's the only way that you could write American economic history, but I had, after all, lived a long time. I had gone through a period when economic history was meant to explain why the American economy was "number one" (as Nixon would say), and, by the end of the 1970s, the question was, "What was it about the American economy that caused it to be something else?" So it seemed to me it was time to examine issues which had been ignored in earlier treatments of American economic history. It wasn't just growth and success we were interested in. It was failure, it was intractable problems, it was decline as well as expansion. We had a lot of things to try to understand, and I'd kept up on all this material.

Do you think you learned anything important just by writing the text?

I learned just how hard it was to keep mistakes out of something that complicated.

I'd like you to talk a little bit about the personal origins of your interest in economic history. How did it happen?

Well, it was very simple for me. As a graduate student I studied more or less the same things the other students studied. I had no more courses in economic history than I had in the history of economic thought or economic theory. I had far more courses in economic theory than I had in history. The two people who had taught me were Morrie and Doug, so that the economic history I had was very vivid in my head. When I got to Oxford I was simply assigned by J. R. Hicks to be an economic historian. He assigned me to be the economic historian; he assigned me to Habakkuk and that was that. Then when I came back to the United States, I didn't do economic history, I went into finance in New York. I was hired at Purdue as a money and banking guy, along with Ed Ames, right out of the Federal Reserve System. But as fate would have it, I ran into Lance again and that was a fateful mixture because, once we got together again, we began discussing the issues that had come up in seminars at the University of Washington. We were joined by Duncan McDougall. We wrote a textbook, and, at that point, we found ourselves faced with the requirement that we teach economic history. It hadn't been taught at Purdue before. When we asked the man responsible for this (why was it we were being forced to teach economic history?), he said the rule at Purdue was that, if you wrote a book on a subject, you had to teach it — God's truth. We introduced American economic history because we had to. So we found ourselves teaching economic history, which we had never intended to teach. It was a lot more difficult to teach than other parts of economics, as you know, and we didn't welcome this new chore.

So you taught it as a team?

No, we would teach it one at a time. We didn't need to be a team because we wrote the book together. McDougall had come there as a macroeconomist, one of the people out of the Kuznets barn. One thing led to another, and pretty soon we found ourselves making reputations for ourselves as economic historians and nothing else. It happened very fast. We had no choice. There was a lot of demand for it. Then came Stanley Reiter and "his word," and the

Cliometric Society appeared. We had that first meeting and found the people who were doing this kind of work. Within just 2 or 3 years we had a huge development coming out of something which had started without premeditation, totally without premeditation.

I've told the story elsewhere about the paper I did with Lance about the exchange rates. That was due to old Arthur Cole. He had seen some of the stuff coming out of Purdue. He approached Lance and said, "I've got the records of these 4,000 bills of exchange. Why don't you and your friends at Purdue see what you can make of these?" So they sent us the records of this company in Philadelphia which they had in the Baker library at Harvard. We sat down across the table from each other and started coding them onto punchcards. That was how we spent days, weeks, months.

If you were putting together a list of people doing significant work on economic history in 1960, how long would the list be?

At that time you had to divide it up between the older people and the younger people, there were very few in between — a topic I took up in the "Fact and Theory" paper. The distinguished senior colleagues, Hal Williamson and Arthur Cole and others, had sort of matured in their subject. Ralph Hidy, Richard Overton, Fritz Redlich, they had a very nice thing they had made that went back to the 1920s, a particular style of economic history which was very mature and was getting to the point where you could use the word historians often use, "magisterial," to describe the output. What then came along was a younger bunch of guys who had been trained in economics whose work could never have been called magisterial and couldn't be to this day because it's a different style. In economics a new theory or a new paper is merely an invitation to an argument. So there is nothing that's finished. No work is finished. I remember one time listening to a discussion about the amount of Cliometrics which had been achieved over the dead bodies of papers that had been proved wrong in the previous meetings. Peter Temin, I

think, asked if there wasn't really some virtue in being right the first time? I think that's a big difference between working in economics, economic theory, and quantitative work and other kinds of production of truth which are based on the quality of the argument and how elegant it is. In Cliometrics you no sooner get finished with something than a new bunch of numbers is turned up that may put you in your place, or show that the theory could be improved by dropping a couple of assumptions which would leave your work standing off to one side.

Do economic historians fare better at that than other economists? Does their work live longer?

No, I don't think so. Well, some of it's going to live. Some of the great errors caused some important new departures; important contributions have been made because of big errors. The errors themselves were important. I don't think economic history is any different from economics in this regard. The half-life of most research is not long.

Does it make sense for economic history to be in an economics department necessarily? The British do it a little bit differently.

Yes, well, I've been of two minds. When things are going well, I think it's wonderful to be in the economics department, and, when things are going badly, I think why should we bother with these people. We've got huge enrollments and could have our own department. I think that ideally we should be in both worlds, in history and economics, and some people like Joel Mokyr, for example, have done it. It's very hard to do because the requirements are quite different for superior scholars in economics and in history. There have been people before — David Landes, for example — who have operated always in both economics and in history departments. There have been some, I think, more successful than others. I think that economic history, because it is now mostly Cliometrics, is so thoroughly involved in computery that it must remain either connected to economics departments or

to mathematical statistics departments. It wouldn't make much difference, you could do it either way, but you should be somewhere where you know when things change.

It's very easy, as I've discovered already, to become comfortable with your own software and your machine and resist. Old people don't like to change, and tend to resist any further changes in thinking. It is the case that the software used and the hardware used also change the way you think. I was appalled by that when I first came back from Oxford and was shown a roomful of IBM machines at Harvard which had been given by IBM to the Economics Department. The department was trying to think up Ph.D. dissertations that would use those machines. I was just appalled by this because I thought the science should determine the technology used, and not the other way around. But you see that when the technology is so sensational, you can't avoid having some of the technology determine the nature of the science. It's just too bad, maybe, but you can't avoid it.

Did you foresee the direction economic history would take after the Cliometric revolution?

No, because there was no way to know what the new advances in hardware, software and theory were going to be. I don't know that I'm surprised by the way it went or even disappointed by the way it went. I alluded earlier to the lack, so far, of big new thinking from the younger people, but then they're younger people. You know, even Eric Jones once worked on small problems. I couldn't have seen 30 years ago how things would be by 1991, obviously no one could. To give you an idea of the extent to which I didn't foresee the future, what I was afraid of, especially after hearing Fogel the first time, was that we might go the route of Oswald Spengler and start developing great algebraic systems of economic history and development that would be bogus. Not that Bob's work was bogus, but it was easy to see how you could go into mechanistic, big systems of, say, difference equations in which you develop mechanistic economic history, and it would be very convincing. On its own terms, it would be irrefutable and would

take up time in classrooms beautifully. It would be very popular because the teachers would only have to learn the sets of equations that were being used.

So why didn't it happen?

Rostow once said that the problem about history was that in the end you were stuck with the facts as they are. Truth is that mankind cannot change, so the tendency to try to explain Western civilization, say, in terms of Kondratieff cycles, or whatever, will always fail because of the chorus of laughter from those who know better. I think some other fields are not subject to that discipline. If anybody says that economic history is not scientific, my answer is: Just publish a mistake sometime and see how long it takes to get caught.

Do you think economic historians, as a group, get along better with each other than other disciplines of economics and are more critical of each other openly?

Well, I think of it with a certain sense of humility. After all, a good education tells you the extent of what it is you don't know. I think economic historians have a lively sense of what they don't know about all the tools they use and don't use, and all the areas they want to investigate and are excluded from, for one reason or another. But it was pointed out, back in the Purdue days, that all the best parties took place when the economic historians were in town. This was not pointed out to me by an economic historian, but by a theorist who was very disappointed about a particular group of theorists who seemed to have no fun at all when they came to Lafayette. When the economic historians came to town, it was New Year's Eve for several days. So I think it was on the basis of that observation you could argue that there is some truth to the idea that economic historians tended to be very congenial with each other.

Another thing you notice among economic historians is that there are not many prima donnas. Prima donnism doesn't pay. I remember several economic historians who began as prima donnas, but it

didn't last long. The first time somebody is able to show demonstrably how wrong you are on a particular point, that really blunts the ego and pulls it down. The thing about Cliometrics, as I pointed out before, the terrifying thing about it, was that for the first time a person doing history could be exactly wrong, could be extensively exactly wrong. You could always make little mistakes, but with Cliometrics, when you set up the capital equipment, that is to say the equations you were going to use to feed those numbers through, if you made a mistake then, you would be wrong for a long time series and a whole analysis based on it. So that kept the egos in check.

I was going to ask you to talk a little bit about "the time of troubles" when *Time on the Cross* was released.

Well, I don't mind talking about that. I was given the responsibility to compare Fogel and Engerman's estimates of slave wages with real wages in England and Europe, which I did at the famous meeting in Rochester. I had studied Grey's *History of Southern Agriculture* under Doug North's fescue at the University of Washington, so it was not news to me that slavery was extremely profitable for those who owned slaves. I was, in fact, astonished there were those who had swallowed the Marxian argument that it wasn't profitable, because it meant you believed a million or so entrepreneurs in the South didn't know what they were doing. The fact that *Time on the Cross* came out in the middle of the civil rights movement didn't add to tranquility at those meetings because a large part of what was being said was for emotional reasons based upon various private agendas about the civil rights movement. Had *Time on the Cross* come out at a time when civil rights were not such an issue, there might have been more media light than there was heat. There was a feeling back in the early 1970s that it was almost illegal to apply neoclassical theory to something that was so intensely human and emotional. Then there were the personalities of the contestants. When Conrad and Meyer came out there had been no such emotional outburst.

But the claims were much narrower.

The claims were narrower. *Time on the Cross* was really provocative. If you go back and look at the research on slavery before 1972, then look at it afterwards, you can see the seminal contribution of Fogel and Engerman, whose work caused a total reassessment of the antebellum South and of the North too. People began to ask "Keynesian" questions about what was so profitable in the South? Were there income effects elsewhere, was an economic rent being collected or what? Then it turned out that the English had cleaned up because of slavery, the North benefitted because of slavery, and the stain continued to spread. That was the contribution of *Time on the Cross*. There were a certain number of hot-headed things said at various meetings at that time, and I thought Bob Fogel and Stanley Engerman were remarkably thick-skinned to go through that kind of fire fight. I'd rather be anonymous than have to go through that. I think there is no long term enmity because of it. Well, I mean those who made successful arguments against Fogel and Engerman then were up-ended themselves. Which goes back to what I said earlier, that it was impossible to be magisterial where everyone had the weapons and were anxious to use them.

And when you write down specific equations you can't go back and say that wasn't quite the number that I intended.

Thomas Edison once said that he wouldn't patent things because a patent was an invitation to a lawsuit. I think the same thing is true of quantitative economic history. Just as soon as you put a number down or write down an equation or draw a diagram, somebody else is going to be able to advance his own position in life because of it. So you must not suppose that you can get away with anything. And after awhile, if you get bruised enough on particular issues, it makes you a very human person.

How would you describe Doug North's influence on your work?

You know of all of the things that I did, the least original was my thesis, and that was done in a certain way for a certain purpose. After that, after I got my degree and got out of England, and then, in turn, left the Federal Reserve System, I never had any interest, and still don't, in doing anything that somebody else has done just because I see the way to stretch it. I think I got that from Doug North. He used to argue that there's no point in being an economic historian if all you want to do is have small ideas. Economic

history is the place where you can have really big ideas. So why get into economic history and not have big ideas? Not explore big things? And God knows that's certainly the way he has lived.

It seems like a lot of students have been influenced by his energy.

Oh yeah. To be a Doug North student is a very suspicious thing to be.

REFERENCES:

[Below we include citations to articles by Hughes and to works of others receiving mention in the text of the interview. For a complete bibliography of Hughes' writings, see Cain, Reiter, and Uselding.] [eds.]

Hughes, J.R.T., "Fact and Theory in Economic History," *Explorations in Entrepreneurial History*, 2nd. ser. 3:2 (1966); reprinted in *Purdue Faculty Papers* and in Andreano.

, "The First 1,945 British Steamships," *American Statistical Journal* 53:2 (June 1958), with Stanley Reiter; reprinted in *Purdue Faculty Papers*.

, "A Dollar-Sterling Exchange, 1803-1895," *Economic History Review* 2nd. ser. 13:1 (August 1960), with Lance E. Davis; reprinted in *Purdue Faculty Papers*.

Komlos, John, "The Height and Weight of West Point Cadets: Dietary Changes in Antebellum America," *Journal of Economic History* 47:4 (December 1987).

Fogel, Robert W., "Nutrition and the Decline in Mortality since 1700: Some Preliminary Findings," in Stanley L. Engerman and Robert E. Gallman, *Long-Term Factors in American Economic Growth* (Chicago, 1986).

Mokyr, Joel, *The Lever of Riches: Technological Creativity and Economic Progress* (Oxford, 1990), ch. 11; also, "Was there a British Industrial Evolution?," in Idem (ed.) *The Vital One*.

Hughes, J.R.T., "Foreign Trade and Balanced Growth: The Historical Framework," *American Economic Review, Papers and Proceedings* 49:2 (May 1959); reprinted in *Purdue Faculty Papers*.

, "Preface to Modelling the Regulated Economy," special supplement on Social Choice, *Hofstra University Law Review* (Summer 1981), with Stanley Reiter.

, "The Great Strike at Nushagak Station, 1951: Institutional Gridlock," *Journal of Economic History* 42:1 (March 1982) [EHA Presidential Address].

Andreano, Ralph, ed. *The New Economic History: Recent Papers on Methodology*. New York: Wiley, 1970.

Cain, Louis, with Stanley Reiter and Paul Uselding, "The Vital Jonathan R.T. Hughes," in Mokyr (ed.), *The Vital One*.

Mokyr, Joel, ed., *The Vital One: Essays in Honor of Jonathan R.T. Hughes*. Greenwich, CT: JAI Press, 1991. Supplement 6 to *Research in Economic History*.

Purdue Faculty Papers in Economic History, 1956-1966. Homewood, IL: Irwin, 1967.

Australian Conference

(continued from page 8)

to take account of historical experience. Frost pointed out that very high growth rates for large cities are not unprecedented, that in the past rapid growth rates have not necessarily produced a deterioration in the physical conditions of cities, and that some of the most environmentally disadvantaged cities have been among the slowest growing. Even when slower urban growth allows per capita income in cities to increase, the physical fabric of the urban environment will not necessarily be improved. The actual development path depends on past decisions relating to land tenure, housing density and urban administrative structures. John Gagg (ANU) thought that the comparative approach taken by Frost, comparing different cities in different countries at different times, should be matched by more detailed studies of the same city over time in order to limit the range of exogenous influences on city growth. David suggested that in one respect historical examples were not relevant to contemporary urban problems, because modern cities are not net killers, and so modern urban growth occurs *both* endogenously and through migration.

Strong parallels between the past and the present were drawn by David Pope (IAS, ANU) in a paper on 'Bank deregulation yesterday and today: lessons of history.' The 1880s and the 1980s, both periods of 'free banking' in Australia, were also periods of mounting bad debts for the banking sector. In 1893 half of Australia's trading banks failed, and although it is improbable that the central bank would permit any major bank to fail in the 1990s, the scale of current bad debt problems prompts asking why the lessons of the past have been forgotten. Pope showed that the deregulation of the financial sector in the 1980s, which was supposed to increase competition and efficiency, had very little impact on the retail banking sector in fact. This, he argued, was entirely predictable from the nineteenth century experience when deregulated commercial banks pursued market share through non-price oligopolistic rivalry, especially by means of excessive branching. A second

aspect of non-price oligopolistic competition was the deliberate acceptance by banks of higher credit risk in a quest for expanded market share, together with an inadvertent rise in credit risk as bank expansion outstripped the supply of skilled managers who had experience in risk assessment. A better knowledge of the economic history of Australian financial markets could have prevented some of the problems produced by banking deregulation in the 1980s. There was some discussion from Jackson and Gregory about the interpretation of the bank cost functions estimated by Pope, which suggested that the spread of cost structures between the banks was both very large and stable over time.

In a paper on 'Anti-business culture and the changing business establishment in the UK and elsewhere', Leslie Hannah (London School of Economics) addressed the argument that Britain's relative economic decline over the last century has been a consequence of deeply ingrained anti-business attitudes. He examined the performance of the British school and university system and found that the elite 'public' schools do not appear to have deterred their pupils from entering business, and that industry and commerce have captured nearly all the rise in number of first class graduates which resulted from the expansion of the universities in the 1960s. These educational changes have now worked through to senior management level: in 1979 the majority of the chairmen of Britain's top 50 companies had been educated in elite 'public' schools, but by 1989 the majority had been educated in state schools and a greater proportion of them had a university degree. He concluded that long-run change in the educational system and business recruitment has made British business much more meritocratic. Sam Williamson (Miami, Ohio) thought that analysis of the top 50 companies could be refined by taking account of changes in the composition of the sample, and Steve Jones (Auckland) and Peter Tull (Murdoch) both wanted more information about non-university and on-the-job forms of training.

Paul Johnson (ANU and London School of Economics) presented a paper on 'The employment and

retirement of older men in England and Wales 1881-1981'. Although census data show an almost continuous decline from 75% to 9% in participation rates for men over 65, adjusting the data for changes in the size of 27 occupational sectors shows that participation rates were almost constant between 1881 and 1921. This conclusion parallels that for the U.S. found by Ransom and Sutch, but the participation trends in England and Wales do not seem to be closely related either to the availability of social security pensions or to any late 19th or early 20th century adoption of life-cycle savings. Williamson noted that the index of occupational concentration used by Johnson to examine changes in sectoral age-participation rates needs to be interpreted with care because it is unclear whether the numerator or denominator drives the index. He also pointed out that since rising retirement rates were not simply a function of rising income, the paper shows that decreasing social security benefits is not an appropriate response to current calls to increase participation rates among older people.

In the first of three papers on industry and business, Steve Nicholas (ANU and University of New South Wales) spoke on 'The new business history: theory, quantification and institutional change,' which presented an agent-principal model of inter-firm and intra-firm organisational design. Using the case study approach, the model was first applied to the design of incentive contracts by the Hudson's Bay Company to control agents in Canada during the 18th century. The second part of the paper estimated a logit model using a sample of 448 British late 19th-century multinational firms to explain the choice between overseas sales offices and production branches. In his comments Hannah warned that the use of theory could close off as much as open up research areas in business history, but complimented Nicholas on his cautious use of his empirical data to test the expansion of British multinationals abroad. Answering Hannah's challenge to map out the next step forward in quantitative business history, Nicholas thought that the use of formal mathematical models which could be calibrated and simulated offered one exciting new direction.

Diane Hutchinson (Sydney), in a paper on 'Organisational costs: the role of corporate culture,' presented a detailed case study of the diversification behaviour of a major Australian cement producer. She showed that some elements of post-merger product diversification can be explained in terms of a costly policy of labour retention which preserved the reputation of the firm as a 'good employer' and enabled it to retain firm-specific human capital and preserve long-term labour contracts. Hannah thought the reputational explanation was plausible, but wondered whether an equally convincing story could be told by viewing the firm not as a network of implicit contracts but instead as a transactions cost-minimizing unit or as a multi-player bargaining game. Jones wanted more information about the demand side, in order to see how exogenous factors affected internal company organisation. David, in a comment on both Nicholas and Hutchinson, suggested that path dependency had much to offer the business historian tracing the evolution of internal organisation.

The last of the empirical studies was by Greg Whitwell (Melbourne) on 'Regulation and deregulation of the Australian wheat industry: insights from economic history into the "great wheat debates".' The orderly marketing of wheat between the establishment of the Wheat Board in 1939 and its demise in 1989 was a success for farmers, but was increasingly seen as a failure by economists who judged it according to distributional efficiency. Whitwell pointed out that the criteria by which the Wheat Board was judged changed over time, and that if economic historians are to understand the historical evolution of the Board they must be sensitive to the many non-economic criteria of assessment used by growers, administrators and politicians. Mac Boot (ANU) noted that although the historical assessment of the Wheat Board seems far more complex than an economic analysis of the wheat market can allow for, this complexity was to some extent a deliberate device used by the wheat industry to resist market forces. Historical detail should not be allowed to crowd out reference to the underlying momentum of economic pressures.

In the final session of the conference, four 'new entrants' to economic history, all recent Ph.D. students, presented short papers on the way in which they see economics and history interacting. In 'Economic history: a client state?', Marnie Haig-Muir (Deakin) examined the epistemological tension that exists between the positivist tradition of economics and the ideographic tradition of history, and suggested that a true synthesis between history and economics will demand greater methodological and epistemological awareness by all parties. Peter Burns (Newcastle), in discussing 'Economic history and economics,' suggested that economic history can take an active role in prodding economics in a dynamic direction, a view also held by Philippa Mein Smith (Flinders) in her paper on 'The value of a Janus-faced discipline'. She particularly stressed the value of economic history in undergraduate teaching as a means of making economics students more alert to the limits of economic analysis and to the way in which the past has informed the present, and of introducing history students to the analytical rigour of economics. Deborah Oxley (Melbourne) examined 'Feminism and economic history' and argued that economic historians and economists should heed the feminist criticism that studies of national income accounts, of resource distribution within and beyond the household, and of human capital formation, cannot properly address these issues until they take full account of the economic role of women. In a broad-ranging comment Nicholas suggested that all four papers over-emphasized the degree to which research in economics is narrow and static, and in history broad and dynamic. He also thought there was a danger in trying to justify economic history only in terms of how it could serve either economics or history, and suggested it can be adequately justified on its own terms as a rigorous academic subject. This prompted a more general discussion about the future of economic history and of long-run economic analysis in Australasia.

Hawke thought it was important to maintain economic history as a subject widely taught in universities, because of the useful frame of reference it gave to students going on to jobs in the commercial and

public sector. David pointed out that this did not necessarily mean maintaining separate economic history departments, something which Australia and New Zealand have inherited from the British university structure, but which is mostly alien to North America. Others, however, felt that without separate departments economic historians would not maintain a sufficient critical mass either to develop their own work or to influence other disciplines. It was agreed that more dialogue between economic historians and economists would be mutually beneficial, and that economics could gain greatly from an appreciation of the way long-run factors affect current and future economic behaviour. David pointed out that the interchange could not be all one-way, and that historians must be prepared to learn more economic theory.

Labor Statistics Conference

(continued from page 10)

Hanes was also asked whether the constant terms for immigrants and nationals were so different because they were systematically going into different occupations. Several discussants also suggested that "years with employer" might pick up some human capital factors, but there was still some concern about the possible bias of workers who moved into or out of covered industries, especially if the propensity to become self-employed differed between nationals and immigrants. It was also hypothesized that the coefficients would all be very different if "children of immigrants" were considered separately. The session ended with an observation that California might be different because "losers" typically moved West, and this was their last hope.

After a long lunch break, a quick rush to the espresso house, and some signing of autographs, the afternoon session began with a discussion of Robert Whaples's (Wake Forest) paper on marriage premia in the late nineteenth century. Whaples and others speculated positively when asked whether the same result would have been obtained if "family premia" had been analyzed instead. An explanation for why women had positive premia (instead of negative) was more

complicated. This issue generated a suggestion for incorporating characteristics of the marriage market using community sex ratios to capture the degree of selectivity. It also suggested a potential tie-in to Goldin's work with the Indianapolis sample to illustrate the characteristics of women in the labor force. In general, participants felt that variations by location and sex composition must be addressed, as well as variations by employment and occupational status, and that the analysis really needed more demographic analysis to be complete. The possibility the marriage premia could have taken another form, like regularity of employment rather than dollars per unit time, was also suggested. Concern about the problem of omitted variables was expressed, as well as concern about the imprecise concept of "human capital." Specialization within the family unit should be taken into account as well as the acquisition of skill. Others wondered whether additional factors, like a home ownership variable or a variable to capture regional variations in the average wage, could be tied in to the story of immobility, but at least one discussant cautioned about placing too much emphasis on the burden of immobility since there is plenty of evidence that families *were* moving.

Price Fishback (Arizona) and Shawn Kantor (Arizona) presented evidence that labor markets in the late nineteenth century exhibited many similarities to modern markets; (i) workers with enhanced human capital and workers in unions received higher wages, with women and children earning significantly less than men; and (ii) workers were compensated for some workplace disamenities, like high accident risk and greater risk of lay-offs. The discussion of this paper focused on the latter conclusion, although a question about the interpretation of insurance purchases—is it to protect human capital investment or is it to protect from hazardous employment—illustrated that it may not be possible to isolate the issues completely. And, as another participant pointed out, even if the distinction between life and risk insurance can be made, "days lost to illness" includes both accident and disease. The authors had already done some work on a related problem using the New Jersey sample where illness was separated

into job and non-job categories. But still the difficult question remained: couldn't the positive coefficients on membership in a benefit society, days lost to sickness, unpaid vacations, and insurance just come from the fact that higher wage earners are better able to afford these amenities? Finally, several participants noted it would be helpful to look at days lost by occupation, firm size, and industry, and to provide more elaboration about the interrelationships among the factors discussed.

The last paper of the day, presented by Jeremy Atack (Illinois) and Fred Bateman (Indiana), used data from the California survey of 1892 to examine the relationship between ill-health and worker-fatigue induced by long hours of work in adverse conditions. Participants asked the authors to offer insight on the following issues: why employers were apparently so ignorant about the relationship between productivity and fatigue; whether workers were really just trying to obtain a higher wage by lobbying for legislation that limited hours; and why workers showed a clear preference for a shorter day on Saturday rather than shorter hours throughout the week. These issues generated a lively group discussion ranging from profit maximization and productivity (including evidence from Indian textile mills) to the trade-off for workers in smaller firms of longer hours for more relaxed discipline policies. One discussant questioned the reliability of the data by wondering whether workers actually reported sick days when they really just needed time off to take care of other business. This was not thought to be a major concern, but the failure of the model to predict changes in health was considered to be problematic. The authors also elaborated on the reasonableness of their age profile result. The discussion ended with suggestions for additional evidence for the "dangerous" occupations that could reaffirm the conclusions. Did they have high turnover rates? Did poor health have any impact on wages and was that effect related to occupation? Did the age profile vary by occupation?

The final day of the conference began with the consideration of Debbie Mullin's (Virginia) paper on the pattern of labor strikes that emerged in the late

nineteenth century. Many members of the group were familiar with related work and offered comments and suggestions for the future. One fundamental point was the need for analysis to cover non-union strikes which, during this period, encompassed a significant portion of all strikes. Several participants also stressed the importance of including a discussion of the changing composition of strikes during this period – many of which involved issues of fairness – and the role that the unions played in this. Quantitative evidence might be available through comparisons of hours of work for union vs. non-union workers, but descriptive accounts of specific strikes would also provide anecdotal evidence. Even though previous work has indicated that the wage effect of unions is small when other factors have been controlled, the absence of collective bargaining agreements during this period suggested to some that the relationship between wage and unions should be stronger than it is now. A distinction between offensive and defensive wage strikes could also provide some insight. It was pointed out that if the employer-union conflict is conceptualized in a game theoretic context, both sides must be modelled. Another participant added that median voter models available in the political science literature might be applied, although subsequent discussion revealed these typically assume contractual agreements. There was also a call for more analysis of strike funds and lock-outs, and as with almost all of the papers, a suggestion to look at the geographical dimension. The session ended on a more philosophical note – thinking of strikes as a means of solidarity itself, even to create a union, and queries as to the motivating factor behind the collection of these data.

The final paper of the conference by Susan Carter (UC, Riverside), Roger Ransom (UC, Riverside) and Richard Sutch (UC, Berkeley) addressed the reliability of the data collected by the various state labor bureaus, with consideration for both methodology, mechanics, and motivation. It was not clear to Sutch that Carroll Wright even intended to provide anonymity for firms, so that the departure from the initial policy of confidentiality likely did not adversely affect the quality of the data. In fact, the authors

concluded that over time, both workers and firms felt more relaxed about the surveys and the information that would result from their collective responses. A question about the basis for gathering certain information from immigrants – namely, were they better off in their home country or in the U.S. – opened the discussion to a much broader question of motivation: why were these data collected at all? What were the advantages to workers and to manufacturers? Carter's insights suggested that workers really welcomed the opportunity to tell their side of the story, while manufacturers wanted to demonstrate that factories weren't bad for people. Other participants joined in with similar perspectives. Another member of the group wondered about the temporal process involved with the emergence of the state bureaus (e.g., at what stage in Kansas industrial development did a bureau come into existence) and what role the survey instrument played in forming the manner in which the participants looked at the world. Specifically, he wondered if the questions that were asked caused firms to keep different records. To the latter question, the authors replied the surveys asked questions about which records were kept; the effects on record-keeping were probably minimal. Since questions about "representativeness of the sample" had been raised during the discussions of many of the conference papers, Sutch reminded the group of different approaches that had been taken in different states. In Maine, the bureau presupposed the characteristics that defined a "typical" worker and then interviewed those that fit the mold; in New Jersey, atypical workers were asked not only about themselves but also about the conditions for their perception of an average worker; and in Michigan, a particular population (e.g. furniture workers) was targeted each year, and the bureau attempted to survey all members of that group. The issue of ideology came up several times. The authors admitted Carroll Wright was not an unbiased individual, but emphasized this should not diminish the fact that he tried to collect unbiased statistics. Another participant underscored this by noting the collection of these data was not an isolated accomplishment – undertaken solely to bring labor and capital together – but was itself part of a much broader intellectual movement.

CALL FOR PAPERS

1992 CLIOMETRICS CONFERENCE

The Thirty Second Annual Cliometrics Conference will be held at the Marcum Conference Center at Miami University May 15-17, 1992. The conference will be hosted by Sam Williamson, Bill Hutchinson and John Lyons.

Support has been received from the National Science Foundation for the 1992 conference, so that we will be able to pay most expenses for the majority of the 50 participants, as we have in recent years.

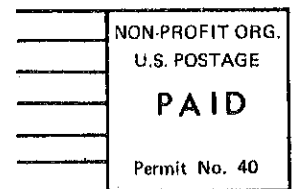
Relevant dates for those wishing to attend the conference are:

Paper proposals and requests for invitations due	February 3, 1992
Notification of acceptance of papers	March 9, 1992
Other invitations mailed	March 16, 1992
Completed 20-page papers due	April 1, 1992
Conference Books mailed	April 27, 1992

All members of the Cliometric Society will receive the call for papers with their December renewal letter. Those wishing to attend the conference should request an invitation by writing to the conference secretary. Those wishing to present papers must send three copies of a three-to-five page proposal by February 3rd.

Cliometrics Conference Secretary
Department of Economics
Miami University
Oxford, OH 45056 USA

THE CLIOMETRIC SOCIETY
DEPARTMENT OF ECONOMICS
MIAMI UNIVERSITY
OXFORD, OH 45056



CLIOMETRICS SESSIONS AT THE 1992 ASSA MEETINGS IN NEW ORLEANS

The Cliometric Society will sponsor four sessions and a cocktail party at the 1992 annual meetings of the Allied Social Science Association. The Saturday session is a joint offering with the American Economic Association. We are trying to organize the Clio Cocktail Party for Saturday, January 4th to begin about 8:30 p.m. Announcements will be made at the Friday and Saturday sessions.

LONGRUN PERSPECTIVES ON CARTELS

Friday, January 3 at 10:15 a.m., Marriott Hotel, La Galerie 4

Presiding: William Hutchinson, Miami University

Peter Z. Grossman (Washington University): *The Express 1851-1913: The Dynamics of a Stable Cartel*

David F. Weiman and Richard C. Levin (Yale): *Competition on the Margin: Southern Bell vs. the Independents in an Emerging Market*

Margaret Levenstein (University of Michigan): *Vertical Integration and Collusion: The Role of Wholesalers in Facilitating Cartel Stability*

Valerie Y. Suslow (University of Michigan): *Cartel Contract Duration: Empirical Evidence from International Cartels*

Discussants: Pascal St. Amour, Queens University

Shane Greenstein, University of Illinois

PROPERTY RIGHTS TO LAND AND RESOURCES

Saturday, January 4 at 8:00 a.m., Marriott Hotel, Mardi Gras A

Presiding: John Nye, Washington University, St. Louis

Sumner J. LaCroix (University of Hawaii at Manoa): *Sheep, Squatters, and the Evolution of Land Rights in Australia: 1787-1847*

Leonard Carlson (Emory University): *Learning to Farm - Indian Land Tenure and Farming Before the Dawes Act*

Ann M. Carlos (University of Colorado) and Frank Lewis (Queen's University, Canada): *Depletion in the Lands of the Hudson's Bay Company, 1700-1770*

Discussants: David Surdam, University of Chicago

Charles Kolstad, University of Illinois

Lee Alston and Don McCloskey were program chairs and have worked diligently to select and group together papers that are sure to hold your interest. We will adhere to our usual Clio style: authors and discussants have 15 minutes to say their piece and then open discussion begins on each paper. Summaries of all papers are attached so you can prepare to join in the discussion.

CONTRACTS & INSTITUTIONS

Saturday, January 4 at 2:30 p.m., Marriott Hotel, Mardi Gras N

Presiding: Trevor Dick, University of Lethbridge

Avner Greif, Paul Milgrom and Barry Weingast (Stanford): *The Merchant Guild as a Nexus of Contracts*

Joseph D. Reid, Jr. (George Mason University) and Michael M. Kurth (Arizona State): *Governments-as-Firms: Occam's Explanation of the Fall of Federal Patronage, 1880-1940*

Werner Troesken (Washington University): *Regulation as Contract: The History of the Illinois Public Utilities Commission*

Ruth Dupré (Ecole des Hautes Etudes Commerciales de Montreal): *"If It's Yellow, It must be Butter": Margarine Regulation in Quebec since 1886*

Discussants: Alex Field, Santa Clara University

Pablo Spiller, University of Illinois

HISTORICAL ISSUES IN LABOR ECONOMICS

Sunday, January 5 at 10:15 a.m., Marriott Hotel, Chartres

Presiding: Samuel Williamson, Miami University

Joseph P. Ferrie (Northwestern): *A Longitudinal Analysis of the Settlement Patterns, Occupational Mobility and Wealth Accumulation of Antebellum European Immigrants to the United States*

Bernard Elbaum and Nirvikar Singh (University of California, Santa Cruz): *The Economic Rationale of Apprenticeship*

Robert A. Margo and T. Aldrich Finegan (Vanderbilt): *The Decline in Black Teenage Labor Force Participation in the South, 1900-1970: The Role of Schooling*

Joshua L. Rosenbloom (University of Kansas) and William A. Sundstrom (Santa Clara University): *Hedonic Wages and Labor Market Integration: The U.S. in the 1890s*

Discussants: Francine Blau, University of Illinois

William Phillips, University of South Carolina

Samuel Williamson, Miami University

SESSION ONE - LONGRUN PERSPECTIVES ON CARTELS
FRIDAY, 10:15 A.M.

THE DYNAMICS OF A STABLE CARTEL:
THE EXPRESS 1851-1913

Peter Z. Grossman
Washington University

The most successful cartel in U.S. business history was almost certainly the railroad express cartel. This cartel, which controlled the parcel post service in the United States, persisted for more than half a century, by one reckoning from 1851 to 1913. During that time, the five firms of the cartel agreed on prices, territory and governance procedures, and, by and large, maintained a harmonious environment within the industry. Stable cartels are, of course, rare historical phenomena. Cartel members typically find it too attractive to violate agreements in order to capture market share. This paper explores the reasons why the express industry proved to be an exception to the general rule.

The express cartel began with a territorial agreement signed in 1851 between Adams and the American express companies to divide the country east of the Mississippi River between them. Over the next decade, three other firms--Wells, Fargo & Co., the United States Express, and the Southern Express--were permitted entry. Thereafter, the five companies apportioned market share. Initially, firms were designated geographic territory, but later they were assigned monopoly control over railroad routes. Where routes overlapped (and members could have competed), the cartel firms fixed prices. Governance procedure were established to adjudicate disputes between members, although in times of crisis, ad hoc bodies might be formed to settle disagreements. Conflicts occurred, but the cartel never broke down. Its eventual demise came at the hands of government that effectively regulated it out of existence in the early part of the twentieth century.

The paper focuses on two key elements in the success of the express cartel--elements that would seem essential for the long-run success of any cartel: 1) its proficiency at deterring (or at least limiting) entry, and 2) its ability to prevent defections among its members. Indeed, in the case of the express cartel, these appear to have been linked. To deter defection, cartel members threatened one another with severe retaliation, typically money-losing price wars. This paper argues that such threats were credible because entry barriers allowed the cartel to function as a monopoly and earn rents. Thus, there was a substantial payoff to long run stability even at the cost of short run losses. To an extent, the express seems to bear out the model of Abreu (1986) who argued that the extent of collusion was determined by the severity of credible punishment strategies. In the case of the express, credibility emerged from the character of the industry.

Entry Deterrence

The express faced potential entry from two kinds of organizations: new express companies and railroads. But the industry had created barriers that for the most part prevented entry by any outside organization. The first, and most significant, barrier resulted from a series of contractual arrangements that allowed the cartel to form a national transshipment system. The system facilitated the shipment of packages throughout the country by any cartel firm. Therefore, any potential entrant faced the prospect of duplicating not only fixed capital, but also a procedural arrangement that lowered transaction costs to

cartel members. Even when an entrant had a production cost advantage, if it were denied access to the transshipment system, the new firm still might have had higher costs overall than the incumbents.

The cartel members also maintained capital sufficient to produce a level of output above the amount demanded of a monopoly producer. This represented a precommitment to remain in the industry, and entrants were left without available profit margins that could easily be captured. An entrant had two alternatives: attempt to grab market share from an existing firm or stay out. The former strategy guaranteed a large expense and a price war.

In models with precommitment, price wars against entrants may be credible by definition. But the history of the express industry provides clear evidence of the employment of such strategies. This was made plain throughout the early history of the express, most notably in 1867 when an outside firm, the Merchants Union Express, attempted what was effectively the last large-scale entry into the express industry. The entrant tried to duplicate a significant portion of the cartel system--largely in the territory controlled by American Express. After spending money to create an express network, the Merchants Union faced an all-out price war. Within a matter of months, the Merchants Union lost the better part of \$20 million--its paid in capital. Although American Express could probably have won if the conflict persisted, the incumbent's own losses were severe enough so that it chose to buy out the Merchants Union with an issuance of new stock. Nevertheless, the entry attempt proved that when threatened incumbents would fight a money losing price war, and could impose the higher costs on the entrant. Even though the original shareholders of the Merchants Union received a buyout, they clearly had made a poor investment; American Express paid only \$0.45 to the dollar. As the express grew richer in the following years, the cartel's wherewithal to fight made a successful general entry even less likely, and none was attempted.

Entry attempts by the railroads posed a different problem. A railroad entry was always local; it could only take over the express service on its own line. But there was a great temptation for rail companies to try it. Because railroads held a cost advantage by control of transportation, operation of the express would appear to have allowed them full monopoly rents instead of the share they received through express contracts. But the contractual network--the transshipment system--made the cost advantage illusory. Denial of access to the transshipment system raised transactions costs sufficiently so that entry was almost always a poor decision by a railroad. Railroads were further discouraged from entry by threats of price wars.

At the same time, the cartel depended on railroad transportation and offered inducements to railroad companies to maintain harmony and discourage entry. Essentially, the standard contract gave the railroads a share of gross express revenues, and the express sought to make the percentage high enough to discourage conflict. Institutional change after 1887 gave additional incentives for railroad-express cartel cooperation. The creation of the ICC forced the railroad companies to reveal, and at times alter, their business practices. The express were initially exempt from ICC oversight. Express revenues, which the railroads received, could be counted on to reflect profit-maximizing behavior instead of regulatory caprice.

Cartel Interaction

In 1880, a railroad pamphleteer portrayed the express cartel as a harmonious organization. This depiction was largely accurate. By a series of contracts, the members fixed rates over common routes and apportioned control over railroad lines to members. The contracts also stipulated governance procedures including bodies to adjudicate and penalize cheating. Because conditions changed, prices and territorial division were frequently revised, but by and large agreements were followed. Though penalties and adjudication procedures existed, there is little evidence that they were ever employed.

This was especially true with respect to prices. Although common routes, and with them the opportunity for price competition, expanded throughout the period, firms seldom deviated from price agreements. In fact, as the opportunity for price competition grew so did the willingness to employ severe punishments to prevent it. The threat of severe retaliation--a money losing price war on a large scale--was the key deterrent. As noted earlier, this was made credible by the success of the cartel's ability to maintain a monopoly. Firms gained monopoly rents. Since price competition suggested that the rents would be dissipated, there was the willingness to accept temporary losses to prevent them.

In fact, the monopoly character of the industry raises a theoretical point. Unlike a standard cartel where the alternative to collusion is competition, the alternative to the express cartel may well have been a single firm monopoly. More generally, if a cartel can maintain successful entry barriers there is no inherent reason why, if it breaks up, a single firm may not be able to sustain those barriers and gain control over the entire market. Price defections then might be seen as an attempt by one firm to drive out the others and gain monopoly control of the market. In that case, a more extreme punishment strategy becomes credible: defection is met by an all-out price war that continues until the defector is forced from the market--a punishment of collective predation. Thus defection is tantamount to suicide, and consequently no one defects; collusion is sustained. While in general the express cartel did not issue such drastic threats, members did so on occasion, and there is evidence that the cartel saw collective predation as a real option. This was especially true of the early years of the cartel. Predation had been used by cartel members against rival express firms, and was threatened in early cartel disputes.

Less drastic price war threats were issued far more frequently and appear to have been taken seriously. As the model would suggest, their credibility was seldom tested since potential defectors knew that cartel members would have felt compelled to retaliate. Agreements became largely self-enforcing; cartel price agreements were generally upheld for fifty years.

Breakdowns, on the rare occasions they occurred, can be traced to exogenous events. (This supports conclusions by Porter in 1983 analysis of a leading rail freight cartel, the Joint Executive Committee.) For example, one period of intracartel strife can be linked to the entry attempt into the express of the New York, Lake Erie and Western Railroad in 1886. But after a period of price warfare, the railroad entry was defeated and cartel price agreements were reinstated.

Territorial disputes between express companies also occurred and, in fact, were more common than price disputes. Territory followed rail lines and thus were continually changing. As lines expanded, monopoly territories would suddenly have competition; overlapping routes would increase. Nevertheless, the cartel was generally able to adjust to changes in the railroad system. Members traded routes to maintain approximate market shares, or on occasion, they shared revenues and expenses on express traffic over certain lines.

Territorial conflict in the early 1890s, however, produced the greatest period of stress of the cartel's history. The United States Express Company sought to gain territorial advantages by offering railroads better contract terms. This effort was largely thwarted by

the rest of the cartel though at a high cost to all firms. Indeed, it produced a unique and exceptionally severe punishment. Cartel members felt compelled to buy up a controlling block of US Express stock. This punishment temporarily ended the firm's autonomy and allowed the cartel to determine the defector's policies from within.

The express cartel succeeded largely because it was credibly able to punish both entrants and internal price defectors. This case poses a challenge for cartel theorists. The standard literature separates entry barriers and cartel stability. In the case of the express, the two were linked and as a result the cartel remained relatively stable for more than half a century.

PREYING FOR MONOPOLY: THE CASE OF SOUTHERN BELL

David Weiman

Richard Levin

Yale University

I. Introduction

Since the publication of John McGee's (1958) provocative paper on the Standard Oil Trust, scholars of antitrust law and industrial organization economists have debated with religious fervor the theoretical rationality and historical prevalence of predatory behavior as a means of securing market power. Using the record of the 1911 Standard Oil case to support his argument, McGee claimed that it is cheaper to acquire a competitor directly than to induce exit or merger by pricing below cost. He rejected the argument that pricing below cost might force a competitor to sell out cheaply, arguing that a price war would partially dissipate the value of the merger. Yamey (1972) responded, without fully rigorous proof, that pricing below cost would nonetheless be a profitable strategy for the monopolist if (1) it lowered sufficiently the price at which the competitor could be acquired or (2) it signalled to competitors in other markets or potential competitors in the primary market the monopolist's determination to eliminate competition, despite the costs of engaging in predatory activity.

Despite recent proofs¹ that Yamey's conjectures hold under conditions of asymmetric information about costs, skepticism remains about whether predatory strategies have actually been employed. Most empirical work on this subject has relied on the interpretation of contemporaneous documents and the descriptions of market participants rather than detailed analysis of quantitative evidence. An interesting exception is the work of Burns (1986), who found that in markets where American Tobacco was alleged to have engaged in price wars, its costs of acquiring competitors were reduced. These costs also varied inversely with a (not entirely convincing) proxy for American Tobacco's reputation.

In this paper we analyze quantitative data and interpret descriptive accounts concerning the response of the Southern Bell Telephone Company (SBT) to competitive entry during the decade following the expiration of the basic Bell telephone patents in 1894. Competition arose rapidly in the south, as it did elsewhere, but it proved less durable than in most other regions. We find that in the period from 1894 to 1900 Southern Bell responded to competitive entry, first, by lowering price or expanding investment in response to the threat of entry, and, second, by pricing below cost when entry actually occurred.

This strategy induced exit only occasionally, apparently because Bell had no real cost advantage over independent telephone companies in small markets. After 1900, Bell focused its strategy on investment in its network of toll lines, which had the effect of isolating the independents against whom Bell continued to fight "wars of attrition." The

combination of pricing local service below cost and controlling access to toll lines forced most independents to sell out or take licenses on terms highly favorable to Bell. By 1904, William S. Allen, AT&T's watchdog over its competition proclaimed victory in Southern Bell's territory: "the opposition in the South seems to be disappearing and there is no evidence that it will spring up again."²

II. The Data

Our principal source of data is an SBT report detailing the "general conditions" of its local exchanges in 1900.³ Edward J. Hall, SBT's president, appended the "Exchanges Report" to a lengthy memorandum to the parent company, spelling out the problems plaguing the anemic operating company and his ambitious and costly plans to remedy them. Constrained financially by wary investors, he explained, SBT could only respond to competitive entry by "slaughtering" its prices and profit margins, rather than providing improved service. If these conditions persisted, he reckoned that the company would just break even, covering its debt service with gross profits.

Although Hall only alluded to the "Exchanges Report," it systematically documents the weak links in the company's system and the impact of competition on its profitability. For each of SBT's 78 exchanges, the report specifies the vintage and condition of SBT's equipment, the volume of business measured by the number of sub-stations, linkages to regional and national toll networks, and the overall quality of service and relations with the public. The report also indicates the net earnings of each exchange during 1899, as well as the value of the capital stock at the end of that year.

The final entry for each exchange is the most interesting and informative. It offers a brief, candid portrait of competitive conditions in the local market. In cases where entry had occurred, the report indicates the size of the independent exchange and the number of "points" or sub-stations served by both companies. In many instances the report also provides an assessment of the equipment and service, toll connections, and future prospects of the independent. Where no competition was present, the report documents any threats of entry or measures taken to thwart it. The former is evidenced by efforts of investors to raise subscriptions or to petition the city council for a charter, while the latter typically involved successful lobbying efforts by SBT to block passage of the requisite ordinances. In several instances, to secure political victory, SBT exchanges had agreed to improve the quality of local service, reduce rates, or construct toll lines to nearby cities.

SBT's network of local exchanges included virtually all cities with a population of at least 10,000. Its coverage of smaller cities and towns, however, was sketchy at best. Bell service reached only 20% of urban centers with a population of 5-10,000, except in Alabama and Georgia where the figure was 80%. Local exchanges varied greatly in size, from 2 substations to over 2000, averaging 338 per exchange. The average exchange represented a capital investment of \$30,000, and earnings in 1899 averaged \$1,350 per exchange. Yet, the earnings of individual exchanges ranged widely from a loss of almost \$10,000 in Richmond to a profit of \$34,000 in Atlanta.

Only 33 (or 42%) of the exchanges operated in "dual" markets, where they competed directly against an independent. This figure seriously understates the extent of competition over local service. An additional 12 (or 15%) exchanges reported a current threat of entry, and 10 exchanges (or 13%) had successfully blocked entry in 1899. The remainder, only 30% of the exchanges, operated in monopoly markets. Competition from independents was strongest in the largest markets and, in general, in the states of the Upper South.⁴

III. The Impact of Competition on SBT's Earnings

Cross-sectional evidence corroborates Hall's claim about the impact of entry on the profitability of local exchange service. The graph in Figure 1 clearly illustrates his dilemma. The graph divides SBT exchanges into ranges representing competitive conditions.⁵ Within each range, exchanges are ranked by the level of earnings. In ninety percent of the markets where local exchanges enjoyed unfettered monopoly, net earnings per station were positive, and they were often very large. Only one-third of the exchanges subject to direct competition, however, reported positive earnings. Where entry had been recently blocked or was currently threatened, exchanges generally earned positive profits, but where entry was currently threatened profits were lower on average than those in unchallenged monopoly markets. An analysis of variance summarizes these results. Exchanges in monopoly markets earned average profits of \$8.67 per sub-station. The threat of entry reduced the earnings by 40% to \$5.25 per station. Finally, and most strikingly, exchanges which competed directly against independents reported average losses of \$3.85 per station.⁶

We develop this finding further through a regression of net earnings per station upon a vector of dummy variables representing the state of competition in each local market, a measure of the independent exchange's market share where competition was present, and a measure of the size of the Bell system to control for possible economies of scale or density. The results conform to our expectations. The estimated coefficients of the market structure variables correctly rank the average profitability of exchanges by competitive conditions. Potential entry more than halved the premium earned by monopoly exchanges, while actual competition forced SBT exchanges to operate in the red. The impact of an actual competitor on net earnings per station was more than 2.5 times that of a potential competitor. Moreover, the more successful were independents at capturing market share, the larger were SBT's losses.⁷ Finally, the profitability of SBT exchange service was positively correlated with size. We have not yet identified the source of these apparent economies of scale, but the result may explain why AT&T concentrated their investments in larger cities.

The evidence on earnings per station from Hall's 1900 report is corroborated by data from two earlier sources.⁸ These documents report the net earnings of each Bell exchange in the years 1896 through 1898, and they report the dates of each competitive entry, and, where relevant, exit. Together, these data strengthen the connection between competition and red ink for individual exchanges and the company as a whole. Between 1896 and 1897 average earnings per exchange fell from \$764 to a mere \$284, or by almost two-thirds. Company profits rebounded sharply in the following year, a point which Hall noted in his cover letter to corporate headquarters. Still, he expressed concern over "complications," presumably arising from the competition. His anxiety, the evidence suggests, was well placed. In the three years covered in these earlier reports, average earnings per exchange averaged \$1500 to \$2200 in markets where Bell enjoyed exclusive rights to phone service. In competitive markets, by contrast, exchanges reported average losses ranging from approximately \$1000 to \$3000.⁹

Without evidence on rates by locality, we can not directly observe SBT's pricing response to actual or potential entry. The lower profitability of exchanges facing competition in 1899 is consistent with three possibilities. One is that SBT may have held prices and service quality constant, lost business to new entrants, and hence failed to cover fixed costs. We can, however, reject this conclusion for the 31 of the 33 competitive exchanges.¹⁰ The remaining alternatives are that SBT slashed prices and margins, as Hall suggested, or incurred large outlays to improve the quality of service. In either case the result would be a reduction in current profits and often large losses.¹¹

Aggregate data on monthly revenues per station, also included in Hall's memorandum, offer some clues about the pricing response to new entry.¹² As shown in Figure 2, nominal rates fell by 50 percent from January 1894 to September 1900, from \$4.80 per station to \$2.40, and real rates declined even more sharply. The decline in telephone rates was neither gradual nor smooth. With the exception of the second quarter of 1895 and the fourth quarter of 1896, the real price declined at an average rate of 0.5% per month during the 1894-1897 period. This trend, however, was punctuated by sharp drops of 19% and 11% in the two exceptional periods mentioned. Subsequently, prices declined at the more rapid average rate of 1% per month. The sharp declines in 1895 and 1896, and the steeper trend thereafter, coincided, not surprisingly, with the first wave of entry into the Southern market. Between 1895 and 1897 SBT faced new competition in 20 local markets, or almost two-thirds of those reporting dual service in 1900. Entry waned in 1899, but gained momentum in the following year, as independents moved into the Lower South.

We explore further the dynamics of Bell's response to competitive entry by combining the data from the 1900 survey with the data on the dates of competitive entry. We expand the regression analysis discussed above by regressing the net earnings per station of each SBT exchange in 1899, as before, on a measure of competitive market share and on a measure of the size of the SBT exchange. Rather than include dummy variables for the state of competition in 1899, however, we use a dummy variable to indicate the year of competitive entry. Although the results are quite robust to variations in the specification, they emerge most sharply when the markets are divided into the following categories: those in which entry occurred prior to 1899, in 1899, and in 1900, and those in which no entry occurs. We also distinguish the two markets in which both entry and exit had occurred by 1899.

The results reveal a dramatic response of net earnings per station to entry. Evaluated at the mean values of SBT system size and independents' market share, the coefficients imply the impact of competition on earnings was most severe in the first year of entry. SBT exchanges facing new entry in 1899 experienced average losses of \$22.87 per station. In markets where entry occurred earlier, earnings recovered from the initial impact of competition, and losses averaged only \$2.58 per station. Exit of competitors restored profits essentially to the levels earned in unchallenged monopoly markets: just \$0.69 short of the \$9.27 per station that was earned on average in monopoly markets. Finally, the regression indicates that credible threats of entry had a substantial effect on SBT profitability; exchanges that experienced entry in 1900 suffered losses of \$0.56 per station in 1899 in contrast to the \$9.27 earned in monopoly markets. An alternative formulation indicates that the adverse impact of threatened entry on 1899 SBT earnings was only half as large in markets where the entry failed to materialize in 1900.

IV. Predatory pricing and acquisition.

Interpreting the cross-sectional evidence dynamically, it suggests a common pattern of response by SBT exchanges to competitive entry into their previously protected markets. Their behavior, we maintain, is consistent with a predatory strategy aimed at restoring monopoly control over local exchange service. In anticipation of entry, SBT exchanges braced themselves by reducing their rates and earning lower profits. In response to actual entry, however, exchanges slashed their prices and earnings, signalling their intention to wage a full-scale price "war." After this initial bout of price cutting, SBT exchanges relented, but still set their prices below cost to weaken their competitors. If successful in driving the independent out of business, SBT regained the market power to restore its prices to levels comparable to those in markets in which its monopoly had never been challenged.

In many cases, however, independents were able to withstand SBT aggressive response, as they enjoyed cost advantages over Bell exchanges, especially in smaller markets. Still, SBT's policy deprived the independents of the cash flow required to finance the expansion and improvement of their equipment. This problem was particularly acute for two reasons. First, lower rates stimulated demand for telephone service, and the unexpected volume of business taxed the capacity of independents. Congestion not only resulted in longer delays in completing calls, thus lowering the quality of service, but eroded the cost differentials which enabled them to meet SBT's prices. Second, the independents relied almost exclusively on retained earnings to finance investment.

Around 1900, Hall conceded that SBT's policy of starving the independents would not work. Price wars drained the company of much needed funds to finance its own investment. Even if the policy was successful in eliminating competition, the costs of fighting could not be readily recovered. Although our regression results do not support his view, Allen claimed that exchanges often found it difficult to raise rates after a price war because of government regulations or potential losses in good will.¹³

More important, SBT discovered that it could not win a war of attrition against the independents with low prices alone. In smaller markets independents operated at lower costs because they adopted "appropriate" and less costly technologies for the provision of local exchange and toll service. In mid-sized markets, such as in the Raleigh-Durham area of North Carolina, independents offered their customers short-range toll service to smaller cities and towns and thus attracted business customers. In both cases, SBT suffered from its parent's corporate strategy of building a national toll system from the top down.

Nonetheless, SBT did not abandon its predatory tactics. Instead, it expanded its strategic repertoire to include preemptive investments in toll lines and the consolidation of its exchange and toll networks through acquisition and sub-licensing. Where necessary SBT continued to resort to low prices. Qualitative evidence suggests that SBT followed a two-pronged strategy. First, SBT undertook an ambitious program of toll line construction to complete its regional long-distance network and strengthen its hold over large urban markets. This investment strategy appears to have been preemptive, intended to increase SBT's market power rather than respond to demand, because at the time SBT reported among the lowest volume of toll service of any Bell operating company.¹⁴

Second, SBT acquired or sub-licensed independent exchanges at nodal locations to block the formation of an alternative local or regional toll network.¹⁵ SBT continued to engage in predatory pricing against recalcitrant independents in those large urban markets that might have anchored an alternative regional network. In the case of Richmond, the evidence suggests, SBT sustained large losses over the period to force the independent to sell out at an unremunerative price, thus deterring future entry. Hall threatened a similar response in Mobile, after the Home Company refused SBT's initial offer.

SBT also sought to acquire or sub-license smaller exchanges that formed local networks. Although sub-license agreements conceded the independent a monopoly over local exchange service, covering a radial distance of five miles, they prevented the independent from connecting to non-Bell lines or constructing competing lines. SBT pursued this policy with a vengeance. Between 1900 and 1906, it acquired at least 82 exchanges and sub-licensed an additional 59 (or 77). In particular, it acquired three local telephone networks in Alabama, Georgia, and the Carolinas. After this aggressive foray, Allen proclaimed the defeat of the independent movement in the South.

¹ See Milgrom and Roberts (1982) for a proof of the second proposition and Saloner (1987) for the first.

VERTICAL INTEGRATION AND COLLUSION THE ROLE OF WHOLESALERS IN FACILITATING CARTEL STABILITY

Margaret Levenstein
University of Michigan

This paper examines, both theoretically and historically, the role of wholesale distributors in facilitating collusion in several nineteenth century industries. Earlier work on collusion in the nineteenth century bromine cartel suggests that wholesale distributors played a critical role in monitoring compliance with output restriction agreements (Levenstein 1989). These firms also invested in reputations as maintainers of collusive prices (or as price cutters in the case of "outsiders" to the cartel). Attempts by manufacturing firms to integrate forward into the distribution of their own products were seen as a direct threat to the cartel. More recent studies of the structure of collusion in the chloroform and bleach markets during the same period suggest that wholesalers played similar, though less formal, roles in facilitation collusion in those markets (Levenstein 1991, Chapter 3).

There are three aspects to the questions raised by these examples, focusing on the structure of the collusive organizations themselves, the nature of competition in these industries, and the organization of the participating firms. First, it is necessary to improve our knowledge of the vertical structure of turn-of-the-century collusive organizations in order to understand the functioning and incentive properties of many of these arrangements. Descriptions of nineteenth century pools frequently refer to common selling agencies being established, or relied upon, but this aspect of pool structure is usually not explored in detail. The research on the cartels in which the Dow Company participated, which provides a more intimate picture of the operating of these cartels than many histories which are forced to rely on anti-trust prosecutions for evidence of cartel activity, suggests that these vertical structures deserve more attention than they have received in the past. The paper reviews the existing literature on nineteenth century pools, focusing on those in the chemical industry (see, for example, Ripley 1916, Eskew 1948, and Dewing 1924). It then analyzes the incentive properties of these organization and compares the more well-known characteristics of joint selling agencies with those of cartels which relied on independent wholesale distribution firms. (Bernheim and Whinston 1985 provides a rare formal treatment of this problem.)

Second, this period was one characterized by frequent charges of "cut throat competition." The most convincing explanation of this rampant price competition is that it was the result of the increased integration of markets (resulting from declines in transportation costs) combined with a high fixed cost (and relatively highly leveraged) structure (Lamoreaux 1985.) This created very impatient firms with no history or pattern of cooperation, and therefore, less ability to refrain from competition. (See Fudenberg 1991 for a review of the literature relating patience, reputation, and the ability to sustain cooperation).

This paper does not challenge that explanation of the increased price competition observed. However, it does raise another explanation of why some industries were less successful during this period than previously at sustaining collusive pricing strategies. This explanation, like the previous one, focuses on the growth of "large" firms, but what is important to this argument is not the increases in horizontal size or fixed costs, but rather the increasing vertical integration of these firms. In many instances, it appears, wholesale distributors were crucial to maintaining earlier collusive arrangements. As a result of vertical integration, these independent monitors of collusive arrangements were displaced,

² AT&T Archives, Box 46 02 01, William S. Allen Papers, Allen-Fish, 1/8/1904. Our interpretation supports the findings of Gabel (1991), who analyzes AT&T's strategy in the Midwest. We present an alternative perspective on SBT's strategy of system-building than the one found in Lipartito (1989a).

³ AT&T Archives, Box 1340, SBT Co., "Report on General Conditions of Exchanges, July 19, 1900;" and Hall-Cochrane, 11/2/1900.

⁴ The chi-square statistic, testing the difference in size distribution of cities with and without dual service, equals 20.18 and is statistically significant. In states below South Carolina competition was limited to the largest cities, whereas in the rest of the territory, it frequently occurred in cities at the lower end of the size distribution.

⁵ Following Hall, we measure profitability by the ratio of net earnings to sub-stations, rather than value of capital invested. The capital stock data are not very reliable. Competitive conditions are ordered as follows: (1) unchallenged monopoly, (2) blocked entry, (3) potential entry, and (4) direct competition. In the subsequent analysis we exclude exchanges with fewer than five sub-stations.

⁶ The t-statistic testing the difference in profitability between monopoly and competitive markets equals 4.60 and is statistically significant at the 1% level. Where entry had been blocked, earnings were not significantly different from those in unchallenged monopoly markets.

⁷ The F-statistic, testing the difference between the coefficient on the actual and potential competition variables, equals 5.86 and is significant at a 5% confidence level. When evaluated at the mean value for dual markets, the market share variable implies an additional loss of almost \$3 per station for SBT exchanges, relative to those incurred at very low levels of competitive market share.

⁸ AT&T Archives, Box 1263, SBT Co., Financing, 1898-1899, Hall-Howe, 5/16/1899, which includes a financial statement and "Statistics of Opposition."

⁹ The differences between monopoly and competitive markets in earnings per exchange were statistically significant at the 5% level in each of the three years from 1896 through 1898.

¹⁰ AT&T Archives, Box 1263, SBT Co., "Statistics of Opposition." In all but two markets, SBT service base grew between the date of entry and 1899.

¹¹ As contemporary accounting methods did not typically distinguish capital expenditures from current expenses, increased investment would immediately depress current net earnings.

¹² AT&T Archives, Box 1340, Hall-Cochrane, 11/2/1900.

¹³ AT&T Archives, Box 46 02 01, William S. Allen Papers, Allen-Fish, 12/27/1902. Allen made this observation in reference to conditions in Parkersburg, West Virginia.

¹⁴ In 1902 the ratio of local to long distance messages in the SBT territory was 1.3%, slightly more than half of the national average; U.S. Census (1902). Of the 33 operating companies SBT ranked 25th in the volume of toll service in 1903; AT&T Archives, Box 1348, Operating Companies, Performance Ratings, 1903.

¹⁵ As president of SBT, Hall pioneered this new strategy; see Lipartito (1989a 1989b). The evidence on SBT's acquisition and sub-licensing policy comes from numerous reports and memorandums in the AT&T Archives; see, for example, AT&T Archives, Box 1340, SBT Co., Sub-Licensees, 1900; Acquisition of Independent Companies in Georgia, South Carolina, W. Virginia, North Carolina, and Alabama, 1902-1911; Acquisition of Richmond Telephone Co., 1902; Acquisition and Sale of Exchanges in North Carolina, 1903; Home Telephone Company of Mobile, Al., Negotiations, 1905. Aggregate statistics on the number of sub-license contracts and the official AT&T policy on sub-licensing are contained in AT&T Archives, Box 1364, Sublicense Statistics, 1907-1908.

and collusion could no longer be successfully enforced. (Of course, if the integrating firm achieved the high market share that it usually aspired to, new methods of achieving monopolistic pricing were available to replace the old, ineffective ones.)

Second, once we recognize the importance of the vertical, as well as the horizontal, nature of these collusive organizations, new questions are raised about our understanding of changes in the organizations of the firms themselves (as opposed to the organization of markets). In contrast to the argument put forth by Chandler (1977) that horizontal and vertical integration were alternative strategies facing firms, we need to understand the intimate connection between the horizontal and vertical changes observed. In the cases of both DuPont and General Motors, we observe simultaneous changes in the horizontal and vertical structure of the industry. These firms were not willing to make the risky investment in vertical integration without the assurance of the modicum of control over market price afforded by the horizontal mergers (and the resulting increased market share) pursued.

Similar concerns dominate the thinking of the Dow Chemical Company in its consideration of vertical integration during this period. However, for the Dow case we have more detailed descriptions of the collusive organizations that existed in the various markets in which Dow participated. We also have explicit discussions of the possible disruptiveness of vertical integration to those collusive arrangements, and the importance of achieving monopoly power concomitant with vertical integration if this was to be a successful strategy.

CARTEL CONTRACT DURATION: EMPIRICAL EVIDENCE FROM INTERNATIONAL CARTELS

Valerie Y. Suslow
University of Michigan

A formal cartel organization sets up a multilateral relationship between firms in a given industry. While standard contracting literature has studied bilateral monopoly relationships in detail, there are horizontal relationships, like cartels, that also have contractual aspects. To shed light on how these contracts work, this paper takes a sample of legal formal cartel contracts and examines their structure and durability. Borrowing from industrial organization theories and transaction cost theories, the empirical model tests for the importance of demand uncertainty and cartel organizational characteristics in determining cartel contract duration. The results show that the more uncertain the environment within which the cartel operates, the shorter the expected cartel contract duration. Other industry structure characteristics and cartel organizational variables included in the empirical model have a much weaker influence on contract duration.

1. Introduction

In his 1985 book, *The Economic Institutions of Capitalism*, Oliver Williamson states: "Many...issues which at the outset appear to lack a contracting aspect turn out, upon scrutiny, to have an implicit contracting quality. (The cartel problem is an example.)" [p.17]. In other words, while standard contracting literature has intently studied bilateral monopoly relationships, there are multilateral relationships, like cartels, that have a contractual quality as well. In the same way that buyers and sellers must transact with each other in a repeated setting, firms in an industry striving to achieve a cooperative outcome also interact repeatedly. But there are important differences. These firms are in a horizontal,

not a vertical, relationship. Cartel members do not transact with each other, but rather, they are connected in their effort to solve the "cartel problem" of intermittent price wars. In order to shed light on how these contracts work, this paper takes a sample of legal formal cartel contracts and examines their structure and durability. The data set assembled for this analysis covers international manufacturing and commodity cartels in 45 industries over the period 1920 to 1939.

Standard industrial organization theory sees cartels as infinitely lived and has focused on those factors affecting occasional price wars.¹ Mainstream empirical literature on cooperative oligopoly has therefore concentrated on whether specific cartels have succeeded in charging a price greater than the competitive price, and if so, for how long.² Any contracting theory of cartel arrangements would absorb many of the equilibrium results of this price war literature. For example, industry structure and demand uncertainty are seen as important determinants of the timing and frequency of breakdowns in pricing discipline. As the environment grows more volatile (or as the number of firms in the cartel grows) the cartel equilibrium point changes and the contract renewal process grows more complex. Also, members are more likely to make a misstep in judging production levels, perhaps triggering a price war, which may in turn terminate the contract.

In the transaction cost literature, Williamson stresses asset specificity, uncertainty, and frequency as the fundamental dimensions along which transactions differ. Although relation-specific investments and transaction frequency play little role in determining the duration of formal cartel organizations, the degree of uncertainty could well be significant. Forms of governance could also affect the longevity of the cartel organization.

Thus, the industrial organization literature and transaction cost literature suggest a list of variables which might influence contract duration. By application of a censored proportional hazards model, this paper examines the empirical relationship between cartel contract duration and aggregate economic uncertainty, while controlling for selected industry characteristics and organizational characteristics of the cartel.

The results show that the more uncertain the environment within which the cartel operates, the shorter the expected cartel contract duration. Both economic downturns and increased economic volatility have destabilizing effects on the cartel contract. Other industry structure characteristics and cartel organizational variables included in the empirical model have a much weaker influence on contract duration. In particular, the number of participants to the agreement has no strong influence on the durability of the contract.

2. The Data

The data set consists of observations on forty-five industries with cartel activity between 1920 and 1939. These manufacturing and commodity cartels operated on an international level. European firms were frequently the organizers and more powerful cartel members. Firms in the United States participated on occasion, by joining 1) illegally, 2) "legally" through the Webb-Pomerene Act, or 3) informally, by having an "understanding" with the organized cartel.³

The overt nature of these pre-WWII cartels sets this sample apart from most previous cross-section samples of cooperative agreements. The major European countries had no systematic antitrust legislation until the mid-1930's. Those few laws that were passed (by Great Britain, for example) took a very lenient stance towards cartel agreements.⁴ Thus, most of these cartels were explicit contractual arrangements (with the occasional exception of U.S. participation) organized in countries which tolerated horizontal cooperation.

Table 2 of the Appendix gives the citations for the data. The data were collected from a variety of secondary sources: books and articles written between the late 1920's and mid-1940's by economists and political scientists, U.S. Congressional hearings, U.S. Tariff Commission Reports, and League of Nations sponsored monographs. Books by Hexner [1945], Stocking and Watkins [1946], and Elliott, et. al. [1946] were used to build the foundation of the data set. Many earlier sources were used to verify and fill in information as needed.

3. Cartel Contract Duration

The duration of a cartel contract is defined by specifying the beginning and ending dates. Cartel contracts begin either on the date the cooperative agreement was signed by participating firms, or the effective date of the agreement when that information was available. Although the sample runs from January 1920 to September 1939, a given cartel contract could begin at any time within the sample period.

A cartel contract ends either when the cartel dissolves endogenously (through the defection of an important member, for example) or because of an exogenous event. It is common when working with economic data sets that some of the observed durations will be incomplete. The outbreak of World War II is the primary reason for a premature ending of a cartel in this sample. These observations are right-censored, and can be accommodated in standard duration analysis.

Determining the approximate ending dates for non-censored contracts is a formidable problem. The question is whether to use the actual contract termination date, as specified in the original contract or subsequent renewal, or to use price and cost data to measure the economic dissolution of the cartel. The first option clearly ignores information. The second option simply is not feasible for this particular sample, nor does it directly address the question of contract duration.

A compromise approach that is feasible is to date the approximate end of an alliance by linking it to some specific event affecting the firms or the market. The process used for determining the ending dates thus relies heavily on a careful reading of statements made by the secondary source documents cited above and listed in the Appendix, Table 2. Cartel contract durations in this study are often measured as shorter than the originally stated contract length. Only for one cartel agreement is it the case that there is no information on the date of termination of the cartel other than the initially stated contract period. In this way the other observables, such as the choice of cartel organization, can be linked to a more meaningful measure of the lifespan of the formal organization.

The data therefore consist of a contract length together with information on whether the contract duration was censored. These contract durations serve as the dependent variable. For censored observations the observed random variable is the contract duration at the time of censoring. Table 1 of the Appendix lists some of the information on the cartels in the sample: the industry or product(s) involved, approximate dates of cartel operation, identities of participating countries, and the apparent reason for the end of the contractual relationship.

4. Empirical Results

The probability that a cartel contract ends this month, conditional on the cartel being intact last month is known as the hazard function. The empirical results summarized here are based on a proportional hazards model where the hazard rate is modelled as a function of time and explanatory variables.

Most industrial organization theories and theories of contracting would predict first and foremost that the larger the number of participants, the harder the organizational

problem. For this sample of international cartels the sign is as predicted, but the magnitude of the effect is much less than anticipated.

There are a few organizational variables that are consistently significantly different from zero. For example, cartels which have formal stipulations for enforcement mechanisms have a lower probability of failure. Cartels built around patent or cross-licensing agreements also have a reduced probability of failure, due most likely to the substantially limited threat of entry. (It may also be that we are seeing the effect of a permanent cross-licensing contract although the price-fixing conspiracy itself may not be.) There are also indications that that agreements attempting to cover a wide variety of products have a lower expected duration. The enforcement problems undoubtedly multiply rapidly as the number of products regulated by the contract increases.

Regarding the cyclical indicators, the empirical evidence indicates an inverse relationship between cartel contract duration and low levels of demand or downward movements in economic activity near the ending date of the cartel. In addition, the data support the hypothesis that both positive and negative "surprises" matter. That is, uncertainty in any direction affects the stable workings of a cartel organization. It is also possible that upswings in economic activity make cooperation (and formal cartel contracts) unnecessary, yielding an inverse relationship between positive volatility and contract duration.⁵

Given the inductive nature of this paper it may be best to turn back to the facts to shed some light on this result. Using the NBER business cycle reference dates for either the U.S., the U.K., or France, roughly sixty percent of the uncensored cartel contracts in this sample ended during a peak-to-trough period. The majority of these ended during the Great Depression. We know that a severe downturn will affect cartel contract duration. And yet, some of the agreements that ended during the early 1930's had survived an earlier downturn. There is no clear cut case.

5. Discussion of Results

The goals of the empirical analysis in this paper are twofold. First, it is a useful exercise to learn more about the structure and organization of cartels as an input to industrial organization theory. Second, an empirical model of the determinants of contract duration can shed light, indirectly, on those factors affecting the costs and benefits of organizing a formal cooperative agreement.

There are two classes of variables of interest, those describing economic uncertainty and those describing the organization of the contract. The empirical results presented lend support to the hypothesis that the more uncertain the environment, the more likely the cartel contract will end. The organizational variables explain much less of the variance in contract duration. While the number of products covered by the agreement and the use of penalties to enforce the agreement seem to matter, other structural characteristics, including the number of participants to the contract, have little effect on the length of formal cartel agreements. The most important contract provision in determining contract length is also the most enforceable, i.e. whether there is a cross-licensing or patent agreement as part of the cooperative contract.

These cartel contracts are conspiracies against the social good. If policymakers want to spur on their demise, the empirical results above imply either that there are no policy instruments or that the only choices high cost. A highly volatile economy will cause cartels to collapse, but minor remedies to industry structure will not. Laws against price fixing may be the most reasonable approach to society's "cartel problem."

References, appendixes, tables and footnotes available at the session.

SESSION TWO - PROPERTY RIGHTS TO LAND AND RESOURCES
SATURDAY, 8:00 A.M.

SHEEP, SQUATTERS, AND THE EVOLUTION OF LAND RIGHTS IN
AUSTRALIA:
1787-1847

Sumner J. La Croix
University of Hawaii

As well might it be attempted to confine Arabs of the desert within a circle drawn of the sands, as to confine the graziers of wool-growers of New South Wales within any bounds that can possibly be assigned to them, and as certainly as the Arabs would be starved so also would the flocks and herds of New South Wales, if they were so confined, and the prosperity of the colony would be at an end.

Sir George Gipps, Governor of New South Wales, December 19, 1840

At the founding of the Botany Bay settlement in 1788, the British colonial government in Australia claimed all lands for the Crown. For the next 60 years the British government experimented with a wide variety of land allocation methods. The government's policy was directed toward concentrating settlement within a relatively small land area and financing immigration from Britain. After 1830 settlers largely ignored the government's policies and raced to begin grazing sheep on the enormous tracts of closed land beyond the official settlement area. This paper's goal is to document and analyze the colonial land policies in Australia with the tools of modern economic theory and to understand how the changing economy then induced changes in land rights. The paper begins by outlining the evolution of land policies from 1787 through 1847. The settlers' rush to occupy restricted land during the 1820s and 1830s is reviewed and the analysis identifies important factors motivating changes in land policies during the 1830s and 1840s and briefly traces the change in government incentives and land policies as sheep farmers became its dominant constituency in the 1840s.

s and briefly traces the change in government incentives and land policies as sheep farmers became its dominant constituency in the 1840s.

I. Land Policies in New South Wales: 1788-1847

Australia's land policies usually emanated from the British Colonial Office, but it was Australia's colonial governors who implemented and often modified the policies to fit the reality of local conditions. It is important to note this "agency" relationship immediately, as many of the policies emanating from London were ill-suited to the conditions of the colony (Burroughs, 1967). Correspondence between Australia's Governors and ministers in the Colonial Office reveals a continuing tension between London's interest in concentrating settlement and Sydney's interest in developing the country. Between 1788 and 1847 there were numerous changes in the land system (Abbott, 1971), and the account provided below highlights only the most important changes.

When the first colonists arrived in Australia, the Governor of New South Wales (NSW) had been given the authority to make land grants to free settlers, emancipists (former convicts) and noncommissioned officers. Land grants usually stipulated that a quit rent of one shilling per 50 acres be paid after five years, that the grantee reside on and improve and cultivate the land, and that timber "deemed fit for Naval Purposes...be reserved for the use of the crown." Governor Macquarie placed a clause in each land grant forbidding the settler to

sell the grant for five years. Settlers usually ignored these restrictions, and the colonial government made little effort to collect quit rents. Governors made only small amounts of land available for settlement, as less than 1,000 square miles (32 miles square) of land had been granted when Governor Macquarie left office in 1821. At the beginning of 1821, Macquarie decided that land grants should be proportional to a settler's capital, subject to a maximum grant of 2,000 acres. Immigrants bringing more than £3,000 to the colony could, however, purchase additional land.

Governor Brisbane's land grants during his four years in office virtually doubled the total amount of land in private hands. In April, 1822 Governor Brisbane instituted a new land policy requiring that the grantee agree to maintain one convict per 100 acres; grants were now restricted by the number of convicts available to be maintained. To expand the amount of land available to settlers Brisbane announced that settlers could purchase (with his permission) up to 4,000 acres at a minimum price of 5s an acre (with superior lots priced at 7s 6d). Governor Darling took office in December, 1825 with instructions from the colonial office to reduce the use of land grants and to move toward a system of land sales. In September, 1826 Darling eliminated the requirement to maintain convicts and instituted a system of land grants to new settlers of moderate means and land sales at fixed prices in lots of 1920 acres. A shortage of land surveyors forced Darling to suspend land sales in November, 1826 and until 1831 grants were virtually the only mechanism for obtaining crown land.

In 1829 the government issued regulations strictly limiting settlement in NSW to the 22,082,200 acres in the Nineteen Counties surrounding Sydney. This prohibition on settlement was followed up with the "Act for Protecting the Crown Lands of this Colony from Encroachment, Intrusion, and Trespass" which allowed the Governor to appoint commissioners of crown lands to enforce crown property rights against squatters.

The experiment with land sales in 1826 foreshadowed major changes in land policy in 1831. The intellectual trigger for the change came in 1829 with the publication of Edward Gibbon Wakefield's tract, *Letter from Sydney* (where he had never visited; the book was written from his cell in Newgate Prison in Britain). Wakefield argued that New South Wales lacked a sufficient labor supply because too many people were allowed to land. He advocated setting a high price on land to restrict land ownership and to increase the pool of labor available to land owners. The state would use the revenue from land sales to subsidize immigration from Britain, thereby relieving the colony's labor shortage. The British Colonial Office adopted many of his ideas when they issued the Ripon Regulations of 1831 which abolished land grants, replacing them with auction sales at a minimum upset price of 5 shillings per acre. Revenues from the sale of crown lands were to be used exclusively to finance immigration from Britain.

After settlers began to cross the Bass Strait from Van Diemen's Land and occupy pasture lands in Port Phillip, Governor Bourke convinced the Secretary of State for the Colonies, Lord Glenelg, to open Port Phillip to settlement and to begin land sales. In July, 1836 the Legislative Council relaxed its official policy of discouraging settlement beyond the boundaries of the Nineteen Counties by allowing individuals occupying these lands to obtain annual licenses. The licenses cost £10 per year. In 1839 the Legislative Council passed a new law imposing a tax of 1d. per annum for every sheep, 3d. for every head of horned cattle and 6d. for every horse, depastured on land beyond the Boundaries in addition to the £10 license fee. In 1838 Lord Glenelg wrote to Governor Gipps ordering him to raise

the upset price on "ordinary" land to 12 shillings per acre. Gipps complied with his order in January, 1839.

In early 1840 the British Land and Immigration Commissioners recommended raising the price of land once again, and the Colonial Office issued a new set of regulations requiring land to be sold at a fixed price of £1 per acre in most areas of the colony. Gipps presented the regulations to the Legislative Council in December, 1840, prompting a flood of protest from landholders and politicians. In August, 1841 the auction system was reinstated, but in early 1843 the colony received the news that Parliament had passed legislation in June, 1842 raising the upset price at auction to £1 throughout the colony. At least 50 percent of the auction proceeds had to be spent on subsidizing immigration from Britain. In 1844 Gipps issued new regulations for grazing sheep on Crown lands which limited sheep stations to twenty square miles, required graziers to renew licences every 12 months, and allowed graziers to purchase land for a homestead. The 1844 regulations were strenuously opposed by the sheep farmers, with opposition being fueled by the depression of the mid-1840s.

Parliament's passage of the Australian Lands Act in August, 1846 established more secure rights for the sheep farmers by providing them with long leases and retaining the high upset price of £1. In 1847 an Order-In-Council provided for the division of NSW into 3 classes of land: settled, intermediate, and unsettled. Leases not exceeding 14 years could be granted by the NSW government to persons occupying unsettled land for more than 12 months. Rent was set at £10 per year with an additional £2 10s paid for each additional 1,000 sheep above 4,000. The Order also granted each sheep farmer a preemption right to purchase large portions of the run at a minimum price of £1 per acre. The leases did not allow the holder to cultivate or sublet the leased lands.

II. Australia's Great Land Rush: 1820-1840

Until 1813 settlers generally played by the property acquisition rules set by the government. This changed in 1813 when three young explorers crossed the Blue Mountains near Sydney and brought back news of a vast plain with scattered timber and adequate rainfall -- in short, lands suitable for sheep grazing. As new immigrants became aware of the enormous expanse of potentially productive pasture land, they set out for the bush to begin grazing sheep on tracts of land as far as 300-400 miles from Sydney. Expansion proceeded in all directions around Sydney, and most settlers obtained the governor's permission to settle the new lands. However, some stockmen pushed south to the Murrumbidgee River and north to the Macleay River without official permission. The "sheep and cattle mania" of 1826 signaled the beginning of a full-blown rush to claim unoccupied crown lands without obtaining a land grant.

In 1824 the government commissioned its first survey of the settled lands. Armed with the results of the survey, the government began gradually to define area of restricted settlement between 1826 and 1829. In 1829 it proclaimed the famous restriction on settlement to the Nineteen Counties, but there is no evidence that the pronouncement did anything to stop the flow of settlers to the bush. The graziers surely realized that the government had no means of enforcing its restrictions, and that the few forces available to the government were more likely to be devoted to maintaining security in Sydney and its vicinity due to the presence of a large convict population.

The transition from a system of land grants to an auction system with a high upset price on land neither stopped the growth of the pastoral industry nor the rush to settle new lands. In fact, it is probable that by making marginal lands in the Nineteen Counties too expensive to settle, the new land regulations actually speeded up the exodus to the unopened lands (Buckley, 1957). By 1839 there were 649 "stations" (sheep farms) outside of the restricted area. Many of the stations occupied vast areas, with stations covering 20-50 square miles being quite common. Throughout the 1820s and 1830s the discovery of new passes through the mountain ranges allowed settlers to scatter over 450 miles from Sydney by 1840. While

most of the best lands in NSW had been occupied by 1840, during the next two decades settlers slowly filled in the less desirable lands.

From less than 30,000 in 1825, the number in NSW increased to approximately 1 million in 1848. In 1844 one-sixth (9,885) of the colony's population and two-thirds of the colony's sheep were located in the squatting districts. Wool shipments to Britain increased from 175,400 lb. in 1821 to 4 million lb. in 1835 to 26 million lb. in 1847 (Roberts, 1968; Kerr, 1962). The rapid growth in the sheep population, the squatting population, and the value of wool exports transformed the squatters from a small group illegally occupying Crown land into a major constituency of the colonial government. When the new land regulations of 1842 set a high upset price on lease, short leases for squatters, and increased station fees, the squatters organized to resist these measures. They hired a lobbyist in London to represent them before Parliament and enlisted the support of the British woolens industry which had been revived by the stimulus of cheap, high quality wool from NSW. With the passage of the Australian Lands Act in 1846, the squatters finally achieved the de jure property rights that had been denied them since the 1829 restrictions on settlement.

The Government's land policies can be categorized into four regimes. In the first regime (1787-1831) governors made land grants to settlers. In the second regime (1831-1836) the government sold land to settlers at auction; land sales were limited due to high upset prices that increased over time. In the third regime (1836-1847) the colonial government legitimized squatter holdings by issuing licenses to sheep farmers occupying crown lands beyond the Nineteen Counties. In the fourth regime (1847-1865) the government maintained high upset prices and awarded long-term leases to holders of sheep runs (farms). In our analysis below we analyze certain themes prominent in each of these periods.

III. Economic Development and Land Policies

The numerous changes in Australian land policy illustrate many important themes running through the modern discussion of property rights formation. Comparisons with U.S. nineteenth-century land policy are instructive, but it is important to remember that Australia was a colony and that the British Colonial Office had its own objectives when it formulated land policy.

A. Concentration of Population and Law and Order

One objective of the Colonial Office's land policies was to prevent dispersion of settlement over a large area. In July, 1834 Governor Bourke appealed to the Colonial office to extend the limits of settlement southward to Twofold Bay. The Secretary of State for Colonies rejected the request:

His Majesty's Government are not prepared to authorize a measure, the consequences of which would be to spread over a still further extent of Territory a Population, which it was the object of the late Land Regulations to concentrate, and to divert for a distant object, not immediately necessary to the prosperity of the Colony, a portion of its Revenues, the whole of which is barely sufficient to maintain in that state of efficiency, which it is so desirable, the various Establishments and Institutions required by the Inhabitants of the Districts already occupied.

The Secretary's rationale for concentration of settlement is very different from those offered by modern economists. Barzel (1989) and Allen (1991) have emphasized that governments have incentives to restrict settlement to limited areas to reduce the cost to government of providing protection to the settlers. The Colonial Office consistently emphasized the importance of maintaining law and order in a colony which had continuously received convict shipments from Great Britain's prisons since its founding. Britain's interest in minimizing subsidies to its prison colony provide a justification for its interest in this matter.

Allen has argued that restricted settlement areas allow a government to meet its "obligation" to provide protection from attack by groups outside the society. He cites the potential for attacks by Indian tribes on settlers at the American frontier as one rationale for the U.S. restricting the scope of allowed settlements. The threat of organized attacks by aboriginal groups in New South Wales was, however, generally much lower than in the United States. While there are records of sporadic attacks on settlers, aboriginals were at a distinct disadvantage in terms of numbers and weaponry. Moreover, attacks by aboriginals would surely have been considered in cost/benefit calculations by prospective settlers. Yet it is interesting to note that after the government provided limited recognition of squatters' rights in 1836, the squatters quickly complained of attacks from aboriginals and demanded additional police protection.

La Croix (1992) has suggested that governments may desire settlement concentration because it reduces the costs of collecting taxes on settlers. This consideration was particularly important in New South Wales, as the original land grant system provided for grantees to pay quit rents to the colonial government. In addition, tax collection is a labor-intensive enterprise, and in land-rich, labor-poor New South Wales, the government had incentives to structure the land system to reduce the amount of labor devoted to collecting taxes.

The Secretary's rationale for settlement concentration, that it is necessary to support public institutions, could be consistent with the tax collection hypothesis or could be based on an externality argument. There is some indication that the Secretary was implicitly making an externality argument:

With concert and mutual assistance, the result of the same labor would probably have been a greater amount of produce; and the cost of transporting it to market would have been a less heavy item in the total cost of production. A different course has, however, been pursued, chiefly, as it appears, owing to the extreme facility of acquiring land, by which every man has been encouraged to become a proprietor, producing what he can by his own unassisted efforts.

An individual farmer would, however, choose a more distant tract of land only if the expected net revenues (adjusted for the higher transportation costs) were higher than the net revenues to be earned by selecting land from within the Nineteen Counties. Any social losses must be due to negative externalities imposed on farmers within the more concentrated area. One possibility is that expenditures on road are now spread over a larger area, and farmers in the concentrated area lose more than farmers beyond settlement limits gain. The government could, of course, act as a first-mover and make expenditures on public goods such that individual farmers are induced to make correct location decisions. Yet local governments are surely influenced by the presence of distant settlers and may make politically efficient and economically inefficient public good decisions. This may explain the intervention of the British Colonial Office with its incentives to minimize expenditures of tax revenue on the colony.

B. Wakefield's Theory, Concentration and Premature Settlement

Wakefield's booklet, Letter from Sydney, was dismissed by the British establishment upon its publication in 1829. Yet during the 1830s Wakefield's theories "won over the Colonial Office" and had enormous influence on Australian land policy (see Philipp, 1960). Wakefield begins by blaming the depression of 1828 (see Butlin, 1986) on the disorganized system of allocating crown land in the colony. He argued that land grants to new immigrants allowed them to become landowners "too soon." Capital could not be productively applied to the land unless labor was freely available. To prevent new immigrants from becoming landowners "too quickly," he advocated selling land at auction at a minimum upset price. Proceeds from land sales would be dedicated to fund subsidizing immigration from Great Britain. The system was self-regulating, for as land sales increased, more labor would be

demand; the proceeds from the additional land sales would then be used to bring more labor to the colony to satisfy the new demands. The critical element in Wakefield's theory is the upset price of land. If the Governor sets the upset price too low, then settlement will be too dispersed and economic development will suffer; and if the Governor sets the upset price too high, it will be a drag on the colony's economic development.

As many commentators (e.g., Burroughs) have previously noted, Wakefield's intellectual influence on British policy was considerable, and the influence of his followers on the Colonial Office between 1831 and 1846 was at least partially responsible for the maintenance of the auction system and its high upset prices.

Wakefield's emphasis on land sales at auction places him in the company of several modern theorists (Haddock, 1986, Anderson & Hill, 1990) who have also expounded on the efficiency of allocated government lands by auction. Anderson & Hill emphasize the inefficiency of requiring investment or use of the land in order to claim property rights to the land. They argue, correctly, that individuals will have an incentive to use the land prematurely (i.e., when production yields negative land rents) if use allows them to establish property rights over future, positive-valued income streams from the land. In the limit such investment would fully dissipate the value of rents from the land. By contrast, sale of property at auction would not occur until the present value of the land becomes positive. While this would also often occur prior to the date when production using the land becomes profitable, there would be no incentive for the purchaser to begin production prematurely.

Perhaps the main difference between Wakefield's model and the models of the modern theorists is Wakefield's insistence on the importance of setting the upset price. By restricting the amount of land available to immigrants, the upset price influences the supply of and demand for labor and keeps the private stock of land from expanding as rapidly as in Anderson & Hill's model. By contrast, in the Anderson and Hill model, the timing of sales has no effect on the discounted value of the revenue as long as the land is placed on the market prior to the time when its rent becomes nonnegative. Presumably the government would charge to the buyer the cost of defining the property rights and recording the transaction. Both models suffer, however, from some defects. If Wakefield's upset price is set too high, the land may not be brought into production even if the annual rent from the land is positive; the lack of sales in NSW after land became available at £1 per acre in the 1840s provides some indication of this.

Anderson & Hill's assertion that the timing of auction sales is irrelevant for the present value of government revenue also suffers from the defect of ignoring the revenue-augmenting role of an upset price. In a thin market without an upset price, the winner of a first-price oral auction pays the value placed on the land by the bidder with the second-highest valuation. A carefully set upset price can increase revenue if it is set above the second-highest valuation and below the highest valuation. Lord Goderich explicitly admitted that revenue was a major consideration behind the adoption of the 1831 regulations in his communication of the new regulations to Governor Bourke. In hindsight, however, it appears as if the upset price was set at too high a level, as sales outside of the cities stagnated during the 1830s and 1840s.

Wakefield's and Anderson & Hill's models both presume that the government has the foresight and the resources to survey its lands prior to offering them at auction and, more fundamentally, that it has discovered that it has productive lands. In New South Wales, government surveyors had difficulties keeping up with the demand for their services. If it was profitable to place most of New South Wales immediately into production, it would be inefficient to force sheep graziers to wait until the land had been surveyed. In this case efficiency would seem to dictate allowing the sheep graziers to occupy the land until it had been surveyed and then to offer it at auction. (Studies of the pastoral industry indicate positive profits from sheep farming if not from wool sales. Of course, profits may be negative initially if farmers have to prepare the land or acquire knowledge about sheep grazing over time.) One problem with this line of argument is that incumbent sheep graziers

acquire "inside" information about the land and would have an advantage at the auction. Individuals who were foolish enough to outbid the grazier would surely be subject to the winner's curse.

C. Income Flows and the Coase Theorem

Field (1990) has recently questioned the application of the Coase Theorem to situations where changes in the initial assignment of property rights generate significantly different flows of income and expenditure. The evolution of land rights in Australia provides an interesting example of Field's analysis. The 1831 Ripon regulations in conjunction with the 1829 restrictions on settlement reflected the decision of the owner of the lands, the Crown, to restrict present use of these lands. If the Crown had been able to enforce its decision, the rapid growth of the sheep industry in NSW would not have occurred. The inability of the NSW government to enforce its property rights (given the enormous expanse of land, the high price of labor to the government, and the growing number and wealth of squatters) led to rapid growth of the pastoral industry. Shipments of wool from NSW to Britain enabled the woolen industry to substitute lower-cost Australian wool for German wool. A revived woolens industry constituted a British interest group that would defend the de facto rights of the sheep graziers to their occupied lands. The changes in land leases and rents in the early 1840s allowed the squatters, a significant economic and political force in Australia after their growth in the 1830s, to join with the woolens industry to place pressure on the colonial office to ensure that the newly established flows of high quality wool would continue. The 1847 Orders-In-Council provided more security to the squatters and ensured that the land would continue in wool production. In sum, the NSW pastoral industry nicely illustrates Field's thesis.

INSTITUTIONS AND AMERICAN INDIAN FARMERS-INDIAN LAND TENURE AND FARMING BEFORE THE DAWES ACT

Leonard Carlson
Emory University

The official policy of the federal government throughout much of the nineteenth and early twentieth centuries was to promote farming by Indians. Both in the 1880s and today, scholars and policy makers were concerned that it would be difficult for Indians to become farmers. Two concerns centered on the possibility that Indian institutions would hinder the success of Indian farmers. The first arose from the belief, shared by most reformers in the 1880s, that Indian tribes did not recognize individual rights to land. Instead, the reformers believed that Indians held land in the form of what economists today call a common, and as long as it was the reformers were convinced that Indians would neither succeed as farmers nor assimilate into white society. Senator Dawes expressed this belief when he concluded about the Five Civilized Tribes that "Till this people will consent to give up their lands, and divide them among their citizens so that each can own the land he cultivates, they will not make much progress."¹

The second concern, often expressed by modern scholars, is that it would have been nearly impossible for Indian institutions to change rapidly enough for Indians to successfully learn to farm in the time frame specified by policy makers. Indeed the consensus of modern scholars is that efforts to teach Indians to farm were fundamentally in error because Indians simply lacked the necessary cultural background to become farmers. Thus Hurt, in the most recent comprehensive survey of Indian farming concludes that ". . . government policy designed to fashion the Indians into subsistence farmers required too great a cultural change

for most of them, even under the best of circumstances."² Hurt also stresses the economic constraints faced by Indians poor lands available to many tribes and the low return to subsistence farming which everyone agrees are very important in explaining the failure of Indian farming. But Indian culture is presented as an important and slow to change factor which made it difficult for Indians to become farmers. Similarly recent a recent paper in the *JEH* also notes in the same spirit that efforts to promote farming ". . . flew in the face of almost all . . . Indian values."³ More broadly, Indians are sometimes cited as an example of a people who have low per capita incomes in part because their cultural attitudes conflict with "the demands of a highly technocratic, rationalistic society."⁴

Not only did the government promote farming, its policy inevitably placed all its resource behind making Indian men into farmers. This was the norm of the household division of labor in Anglo-American society. But in most tribes outside the Southwest, agriculture was women's work. Josephy, for example, concludes that one of the reasons for demoralization among plains Indians was that "The hunters and warriors, stripped of their dignity and self-respect, were given few *manly* diversions, and many of them, *losing the respect of the women and children*, sank into an indolence that withered their souls and turned them, ultimately, to alcohol as an escape and violence for an outlet for their hurts."⁵ (emphasis added) Similarly, Novak concludes that "Sex roles were also strictly defined. . . in general men hunted and women farmed." (p. 641)

Thus farming as promoted by the federal government meant that not only would Indians have to change their way of making a living, but the division of labor within the family as well. It is not surprising that many have concluded that the needed changes were simply too great. In the end, however, whether or not Indian institutions were able to change rapidly enough is an empirical question.

If institutions matter and constrained Indian efforts to use their available resources, we need a clear understanding of which institutions mattered and how if we are understand the interaction of culture and material economic success. This essay extends my previous work to examine how pre-reservation Indian land tenure and other arrangements influenced the success or failure of Indian farming prior to allotment under the Dawes Act. Two main questions are examined: (1) what was the nature of Indian agricultural traditions, especially land tenure, but also the household division of labor, prior to allotment; (2) were Indian institutions sufficiently flexible to allow any tribes who were not already farmers to succeed in the late nineteenth century as subsistence farmers?

Indian Land Tenure Prior to the Formation of Reservations

Because reformers in the 1880s did not believe that Indians had the necessary institutional background for settled agriculture or that Indians were capable of establishing the "proper" set of institutions themselves, reformers in the nineteenth century tried to push Indians to become farmers and at the same time assimilate American values by imposing a modified version of Anglo-American property right on them through the Dawes Act of 1887. The result was a heavily restricted and inefficient form of property rights.⁶ This policies rested on a misreading by the reformers of the nature of existing institutions devised by Indians to limit access to resources and organize their use. In fact neither the reformers or their critics in the 1880s and 1890s had a clear understanding of how Indian land tenure worked prior to allotment. According to D.S. Otis "Friends and enemies of allotment . . . (b)oth were prone to use the word "communism" in a loose sense in describing Indian enterprise. It was in the main an inaccurate term." (Otis, p. 11.)

Today relatively few Indians make their living as farmers and the years following allotment saw a decline in the land base available to Indians. The failure of Indian farmers in the twentieth century is compelling evidence to many that Indian institutions were a major barrier to Indian farming. But the failure of Indian farming after the Dawes Act is

not necessarily mean that Indians could not have become farmers under a different set of circumstances. Since they did not understand the richness of Indian institutions for controlling access to and use of resources, the reformers who were the guiding force behind the passage of the Dawes Act of 1887 simply did not allow Indians to make use of institutions that might have worked on reservations.

Thus it is important to understand the nature of Indian institutions both before and after the creation of reservations and to examine closely how Indians fared as farmers at different times. While there was a wide variety of these institutions, land tenure arrangements can usefully be categorized into a few types. According to Linton ". . . the linkage between ecologies and basic economies in North America is close enough to make a description of land tenure by culture areas fairly valid."⁷ Which is what we would expect where tribes adopted institutions suited to their environment.

In the East and Midwest, many tribes, classified as eastern agriculturalists, had economies which depended on a mixture of hunting, by men, and horticulture, by women. A tribe typically controlled a recognized territory. Within a tribe's territory, fields were cultivated by an individual woman or by a matrilineal kinship group, with a usufructory right being held by the individual family or kinship group.

Because of external military threats, the Iroquois lived in central villages practicing settled agriculture. Fields were farmed by lineage groups of women who resided in the same long house under the direction of the oldest woman, the lineage head. Shirking was constrained by the fact that work was done by women who were related to one another and living together.

Another important eastern group were the so called "Five Civilized Tribes" (Cherokee, Choctaw, Creek, Seminole, and Chickasaw) most of whose members were forcibly moved to Oklahoma from the Southeast in the 1830s. The largest of these tribes was the Cherokee. At the time of their removal the Cherokee had had substantial contact and intermarriage with Europeans. A small group were planters with extensive commercial holdings and slaves, while a large number had a more traditional orientation and generally had small subsistence farms of roughly five to seven acres cultivated by women, while men engaged in animal husbandry and hunting. Using quantitative materials, however, Wishart finds that a significant number of Cherokee who could be classified as "yoeman" -- that is farmers who produced more than a minimum subsistence level.⁸

In the more arid Southwest, land use rules among agricultural tribes were more complex, but there, too, use rights to land were typically recognized by the tribe. The Pueblo and Hopi peoples lived in settled villages with sophisticated irrigation technologies requiring relatively centralized decision making. The Apache and Navajos, on the other hand, were raiders and herders and who had no need for a system of centralized decision making.

For those engaged in settled agriculture, ownership of the land itself rested with the tribe or the clan, but again the use of the land and its improvements was recognized as belonging to the individual or the family. Centralized control of resources or direction of communal labor by tribal leaders existed for the provision or maintenance of public goods such as an irrigation system. Among the Pimas, bringing a new tract of land under cultivation required the labor of all men in the affected villages to build the irrigation system and related improvements. This was done under the supervision of the tribal elders, who also supervised its own going maintenance.⁹ Following the construction of the canal, the village headman, with the aid of an advisory council of community leaders, assigned farm plots to those who had participated in the work. Those farm plots became the inalienable property of the assignee and his heirs although, it could be loaned to others.

Among the Pueblos, land also was used by individuals but it belonged to the tribe. There were obligations to provide labor to support religious leaders and to participate in community labor to help the needy. The Pueblos always recognized individual ownership of

the animals, which grazed upon village lands, although the number of animals an individual could graze was limited.

Among the Hopis farming was carried on by individuals or families on land which belonged to the clan of the wife. Land which did not belong to a clan could be farmed by anyone but without right to will or sell the land. Interestingly, the introduction of peach trees by the Spanish led to a modification in this system because the trees required a longer term investment. The person who planted and cared for orchards could sell or will the trees.

In summary, land tenure arrangements found among native peoples in the Southwest were consistent with the relative scarcity of resources and the costs of enforcing property rights. Rangeland, which was abundant and difficult to enclose, was treated as a common, but improved farm land was scarce and therefore individually owned even when allowed to lie fallow. Economic activities were largely conducted by individuals or families with individual rights to land and animals which were recognized by the tribe. Centralized management of group activities -- which meant monitoring individuals who might shirk -- were common only for the construction and maintenance of public goods such as irrigation systems.

Other tribes, such as those of the Great Plains or Rocky Mountain Plateau, were more nomadic, did not engage in settled agriculture, and did not recognize individual claims to land. By the eighteenth century, Plains tribes had an economic system based on using horses, introduced by the Spanish, to hunt bison. Prior to the eighteenth century, the western plains tribes had either been hunter gatherers while eastern plains tribes had lived in villages practicing mixed hunting and farming. Most bands of the Sioux nation pushed west by the Chippewa completely adopted the lifestyle of a plains tribe and ceased entirely to plant crops. The Santee Sioux, however, continued to live in settled villages. The adaptation of the plains culture in the early eighteenth century illustrates how rapidly and successfully agricultural tribes could adjust culture and institutions to new conditions.

The tribes who lived in the plateau region of the Rocky Mountain states usually did not plant crops and migrated in search of a variety of food sources. nor did they recognize individual land tenure. These tribes had recognized land ownership of territory, but did not recognize individual land tenure. "Towards the south, where food became increasingly scarce, even band lines seem to have broken down and people lived in isolated families. . . Apparently existence was so precarious that all resources had to be shared by all." (Linton, p. 48)

The Indians of California, like those of the Rock Mountains depended upon wild plants, especially acorns, for much of their diet. Villages controlled well defined territories, but since they did not plant crops, these tribes did not recognize individual rights to land.

The primary economic activity of the Indians of the northwest was fishing. Villages were recognized as owning territory, with the village sometimes laying claim to offshore fishing grounds. According to Higgs, "Indian regulation of the fishery, though varying from tribe to tribe, rested on the enforcement of clearly understood property rights. In some cases these rights resided in the tribe as a whole; in other cases in families or individuals; sometimes in a mixture of the two."¹⁰ Planting and hunting were less important among these tribes, and rights to land were not well developed.

The evidence about Indian land tenure prior to tribes being confined to reservations is consistent with the hypothesis that Indian institutions were efficient. Hurt reaches a similar conclusion for agricultural Indian tribes "The Indian concept of land tenure enabled various villages to make the best possible use of the land in order to meet their own specific needs. Each people also developed a rational system for transferring land after the death of the owner" (p.75). In other words, these property rights adjusted to the time and place specific relative prices of each tribe. Tribes without settled agriculture did not develop individual rights to land, again because these were consistent with resource endowments and technologies.

Thus there is agreement that Indian institutions seem in retrospect to have been rich and well suited to solving a variety of economic problems. Looked at this way the view that Indians could not become farmers due to the nature of their institutions rests on the conclusion that these institutions, well suited to their original environment, could not change fast enough to adjust to settled European style agriculture. And certainly all tribes faced a dramatic challenge once they were confined to reservations. The largest changes, of course, were required by those nomadic tribes which lacked a tradition of private property rights in land and whose life style was predicated on constant movement. But even tribes such as the Cherokee which had long agricultural traditions faced challenges. With the expansion of European settlement, hunting, which had been the primary activity of men, was no longer a reliable food source, for example. Nor were they used to growing food for sale in a market.

Indian Land Tenure On the Reservations

Once a tribe was confined to a reservation, a key question for this paper is what form of land tenure was established. On the closed reservations, the system that evolved was one of use rights. Typically, the agent and members of a tribe recognized an individual's title to animals and, where farming was practiced, a family's claim to the land it worked. More land could be added to the holding by bringing the land under cultivation. On reservations where cattle ranching was the preferred activity, cattle were individually owned with grazing land open to each individual. Such a system of use rights is consistent with common sense, a sense of fairness, and the common Indian practice of treating economic activity as individual endeavor. A remarkable and too little appreciated aspect of pre-allotment land tenure is how similar it was to that which existed among agricultural tribes prior to being confined to reservations.

It is often claimed that only the Cherokee or members of the other Five Civilized Tribes were willing to farm. This was true at first, but not later. Most measured output of Indian farmers in the 1870s was produced by members of the Five Civilized Tribes in Oklahoma. Not only did other tribes have much to learn and many changes to make, but initially there was often little reason to farm. Traditional means of earning a living were intact and Indian lands were far from markets and, hence, there was little incentive to change. But the environment in which Indian lived changed rapidly. Once the bison herds were destroyed or other similar traditional food sources eliminated by the Army, trappers, and white settlers, farming often became a more attractive potential occupation for many Indians and many more tired farming.

Indian Farming on the Reservations

Evidence to support the view that Indians who had few agricultural traditions were able to farm once there were sufficient incentives to make this an attractive alternative. are found in the *Annual Reports of the Commissioner of Indian Affairs*. Aggregate Indian output and acres cultivated by tribes other than the Five Civilized Tribes grew rapidly after 1875. Acres cultivated rose from 117,267 in 1875 to 369,974 acres in 1895, and an index of output of grain grew by 5.5% per year. Per capita levels were low but there was indeed growth in output. Another measure of the extent of Indian farming is provided by the *Census of Agriculture for 1900*. The Census reported a total of 19,910 farms for 1900. Of these, 4037 were in the North Central region, inhabited by tribes who were often nomadic or at least recently confined to settled reservations. The median Indian farm produced output 18% as large as that of white farmers in the same region. Thus while it would be incorrect to describe Indians as successful commercial farmers, a number were making progress doing just the sort of subsistence farming advocated by the federal government.

Another feature of the census reports is the fact that most of farmers were men. The census of 1910 found 20,841 Indian men whom it reported as farmers and 23,291 who were farm laborers; while it found only 1,156 women farmers and 3,197 farm laborers. The agents

reports for 1900 similarly fail to mention Indian women who were farmers, although it seems likely they probably would have mentioned it if there very many, if only to criticize the practice. Somewhat later, the Meriam Report of 1928 reports that in the 1920s most Indian men in a major survey (68%) listed their occupation as farmer, while there is no mention at all of farming by women in a major chapter on Indian women's household and industrial jobs. The evidence is that despite cultural barriers, by 1900 farming was largely a male occupation. Of course, it is possible that the government's emphasis on making men into farmers may have been excessive -- leading to an underutilization of the efforts and talents of women -- but it does indicate that a shift in the gender division of labor had already taken place.

The Yankton Reservation in South Dakota provides a useful examples of the evolution of Indian property rights before allotment. The Yankton Reservation was established by treaty with the Yankton Dakota (Sioux) Indians in the late 1850s. After the buffalo vanished from the prairies to the West in the 1860s, the agent increased efforts to promote Indian farming. The first step was the creation of an agency farm, supervised by the agent and hired white farmers. The second step saw the growth of farming by individual Indians in addition to the agency farm. By the time of allotment in 1891, the agency farm had been abandoned. As early as 1878 farming was conducted by "each man to himself on his own plot of ground." (BIA, 1878, p. 47) These plots ranged from 5 to 15 acres each. The agent's report for 1888 indicate that individual claims had been recognized for as long as 20 years. (BIA, 1888, p. 20.)

This is true for both farming and ranching. Examples cited in the paper include the Santee Sioux, the Yankton Sioux, the Coeur d'Alene in Idaho. Some tribes had resources and traditions best suited to ranching and some succeeded as ranchers for a time, despite periodic resistance from the agents.

Not all tribes were successful of course. Some tribes, such as the Uintah and Ouray in Utah, for example, initially had few people who were willing to farm. Culturally, there was opposition to farming by among the Ute Indians who lived on the reservation, since products of the earth were the province of women. But it is also true that the reservation was given to frequent droughts and that those Utes who did farm were often disappointed with the outcome. Thus the failure of Utes to farm cannot be solely attributed to the failure of Utes to respond to incentives to farm.

Conclusions

The conclusion is that Indian land tenure prior to reservations were well suited to the resource constraints faced by tribes and not the uniform system of free access that the reformers believed. Further, many tribes were often able to adjust land tenure and other social institutions to the requirements of small scale subsistence farmers prior to formal allotment. Linton reached a similar conclusion about the flexibility of Indian land tenure, noting in 1942 that "Certainly where individual land tenure has been of obvious advantage to the individual there has been little resistance to it." (p. 54) Further Indians who were members of tribes other than the members of the Five Civilized Tribes produced a sizeable amount of agricultural output by 1900. Thus many Indians were at least trying subsistence farming in the late nineteenth and early twentieth century.

Institutions and how they change has received renewed interest among economists. If we are to conclude that Indian institutions made it difficult for Indians to become farmers, more must be said than Indians had inflexible institutions. Some important institutions were more flexible than is often realized, where this was in the interests of Indians.

1 Cited in Otis, D. S., *The Dawes Act and the Allotment of Indian Lands*. Prucha, F. P. (ed.), (Norman, 1973), 10-11.

2. Hurt, R. D., *Indian Agriculture in America: Prehistory to the Present*, (University of Kansas Press, 1987), 233.

3. Novak, S.J., "The Real Takeover of the BIA: The Preferential Hiring of Indians," *JEH* (Sept. 1990), 641.
4. Kaufman, B. E., *The Economics of Labor Markets*, 2nd ed, 26.
5. Josephy, A. M. Jr., *The Indian Heritage of America*, (New York, 1968), 350.
6. Carlson, L. A., *Indians, Bureaucrats, and Land: The Dawes Act and the Decline of Indian Farming*, (Westport, 1981), and "Land Allotment and the Decline of American Indian Farming," *BEH* (April 1981).
7. Linton, R. M., "Land Tenure in Aboriginal America," in *The Changing Indian*, O. LaFarge (ed.) (Norman, 1942), 44.
8. Wishart, D. "Development of the Cherokee Economy Prior to Removal: the Statistical Record," unpublished.
9. Officer, J. E., "Arid-Lands Agriculture and the Indians of the American Southwest," in *Food, Fiber, and the Arid Lands*, (Tucson, 1971), 58.
10. Higgs, R., "Legally Induced Technical Regress in the Washington Salmon Fishery," *Research in Economic History*, 1982, 60.
- 11 See, for example, North, D. C., *Institutions, Institutional Change and Economic Performance*, (Cambridge, 1990).

DEPLETION IN THE LANDS OF THE HUDSON'S BAY COMPANY: 1700-1763

Ann M. Carlos Frank Lewis
University of Colorado Queen's University

I have yet to see any problem, however complicated, which, when you look at it in the right way, did not become still more complicated. (Indians, Animals, and the Fur Trade, p.61)

That the beaver population was depleted during the fur trade period is generally accepted. Irrespective of which part of the extensive body of literature that one might examine, the basic theme is that the Indians in Canada overexploited the fur resource. Indeed, this belief is so widely held that no one has examined the timing, pattern, or scale of that exploitation. This paper addresses just these issues for the period 1700 to 1763, through an examination of the Hudson's Bay Company's post records of the interaction between trader and Indian. We ask whether the company's records of the numbers of furs brought to the post indicate depletion of the beaver population. If depletion is indicated, we then ask over what period did this occur, under what set of economic circumstances, and in which of the Hudson's Bay Company's post hinterlands?

Although there was a Hudson's Bay Company presence around Hudson Bay from 1670, trade at company posts, at least until 1713, depended on the success or failure of its military campaigns against the French. Because of this, we focus on the years between the Treaty of Utrecht (1713) and the Treaty of Paris (1763), which marked the end of French rule in Canada. During these years, Hudson's Bay Company policy in Canada was passive but vigilant in that the Company waited for the Indians to come to trade at the Bay side posts rather than actively going out in pursuit of the Indians. Such a policy operated to the Company's advantage as long as no impediment cut the flow of Indians to the Bay. Because each river-mouth fort commanded a different region, a roughly radial trading pattern developed. In contrast, the French trade from the St. Lawrence can be thought of as linear. Traders from the St. Lawrence were continually pushing further west in search of cheaper furs.

On a theoretical level, the organization of the two trades was similar in that the Hudson's Bay Company had a technical monopoly of the beaver trade in the drainage basin of Hudson Bay, while the French company, the *Compagnie d'Occident*, obtained in 1718 a monopoly of the beaver trade in the St. Lawrence basin. In other respects, however, the companies operated quite differently. The French monopoly issued *conges* (licenses) or leased out the use of its posts. Both leasing and licensing created an incentive for those involved to try to expand the trade as much as possible, therefore it is not surprising that during the period after 1720, there was an expansion of the French trade. But it was not until the beginning of the 1740s that new French posts were established within the western drainage basin of Hudson Bay. Thus it was only for the period from 1740 until the mid 1750s that these posts represented a threat to the flow of furs to Hudson Bay. From 1755 to the British conquest, French competition was a much less serious problem.

In comparison to the licensing and leasing system of the French company, the Hudson's Bay Company was organized on strictly hierarchical lines and attempted to set and maintain uniform rules for the managers of its various forts. The Company dictated the prices managers could charge for goods traded and for furs purchased. These standards were known as the *Official and Comparative Standards* and were denominated in the Company unit of account - a Made Beaver (MB). However, because the nature of the trade required price flexibility, the Company allowed some differences between its posted schedules and those used in actual trade. This difference was known as the *Overplus*. But even here, the head office demanded a full accounting of the actual trading prices and the amount by which the managers had deviated from the posted schedules in order to ensure that no manager used the difference for personal gain. It is within this highly centralized structure that our discussion of depletion takes place.

The study of the fur trade in Canada has generated a large and varied literature, much of which is qualitative. Yet the absence of quantitative analysis, in some cases, has not been dictated by a lack of available data. The Hudson's Bay Company preserved its trading post records; and these records allow us to form a good picture of some aspects of the fur trade including the economic relationship between the Indians and the Company. Arthur Ray and Donald Freeman in their seminal works on the fur trade have already made use of some of these data. In this paper, we also exploit these records to address a number of questions including the specific issue of depletion. By combining the Company records with recent work on beaver ecology, we attempt to infer the change over time in the beaver population and the impact of French competition on that population. In addition, we use the current economic literature on biological resource extraction to indicate the effect of the Hudson's Bay Company pricing policies on the depletion of the beaver stock.

Our analysis is based on a fisheries model of resource extraction that has been applied successfully to fur-bearing animals. The model assumes that the unexploited population follows a logistic growth curve from which can be derived a quadratic relationship between the animal stock and the absolute rate of natural increase:

$$(1) \quad F(x) = aX - bX^2, \quad a, b > 0$$

where X = the animal population and $F(X)$ is the natural growth of the population per period. In equation (1), parameter a , is interpreted as the maximum proportional rate of growth of the population, a rate that decreases continuously as the population declines; and parameter b , equals a/X where X is the (maximum) population if the resource is not harvested.

The observed population dynamics of an exploited species will depend on the harvest rate and the extent to which that harvest varies over time. Combining equation (1) with the harvest rate, we derive the change in population per period as:

$$(2) \quad \dot{X} = aX - bX^2 - H,$$

where H is the harvest. Whether the population is growing or declining depends on the relationship between natural growth and the size of the harvest. In particular as H approaches aX , interest in some resource analysis is the population that yields the maximum sustainable

harvest, X^m . This is derived by maximizing H with respect to X subject to $= 0$. One objective of the paper is to determine the extent to which the Hudson's Bay Company and the Indians followed policies consistent with a population of X^m .

We estimate the beaver population in three areas of the Hudson bay hinterland; the areas served by Fort Albany (including Moose Factory), York Factory and Fort Churchill. Fort Albany, located on James Bay, was furthest to the southeast and hence faced the most competition from French traders. York Factory, more than 500 miles to the northwest, did not experience significant competition until the 1740s and Fort Churchill, which was even further up the Hudson Bay coast from York Factory faced only modest French competition beginning in the 1750s. The annual number of furs traded at these posts provide the basis of our population estimates.

We derive estimates of the beaver population and the French harvest by combining fur trade data with recent ecological findings on beaver life-histories and the way those histories respond to exploitation. Central to our calculation is the natural growth function, $F(X)$. To derive the parameters of that equation we use the ecological studies which indicate that at maximum sustainable yield, 25 percent of the beaver stock is harvested annually. This is consistent both with the recent observed harvests and with the observation that in an established lodge, 3 adult beaver of the 12 would leave each year to establish new colonies. A harvest of 25 percent at maximum sustained yield implies parameter, a , in equation (1) is 0.5. The value of parameter, a , depends on the maximum population a region can support and therefore will vary by trading area.

Our estimates of and the French harvest are derived by selecting plausible values and then determining if those values generate a population pattern consistent with the observed Hudson's Bay Company harvests. Because of the nature of the dynamic relationship between population and exploitation (equation 2), it turns out that only a small range of values could have generated the harvests that we observe. All others either imply an extinction path or an implausibly large population towards the end of the period when harvests were very low. To illustrate the approach, consider the hinterland served by Fort Albany. We begin by selecting a period when it most appeared that the Hudson's Bay Company was harvesting at a maximum sustained yield basis. That is 1718-24, when an average of just over 17 thousand furs were traded. We then test different levels of French harvests, until we find one that generates a consistent population series. This turns out to be just under 4 thousand furs. Thus we assume 21 thousand was the maximum sustainable harvest in the Fort Albany hinterland. Our estimates of the beaver population and the rate of French exploitation are speculative, but at least they are consistent with the observed Hudson's Bay Company trade, the recent beaver ecology literature, and the qualitative accounts of the fur trade. If we are prepared to accept these simulations as reflecting what actually occurred, some interesting insights and interpretations emerge.

The experience of York Factory, in particular, illustrates the central role of the French in the depletion of beaver stocks. Trading began in 1715, but it was not until 1741 that the French built a post in the hinterland, although they had begun trading several years earlier. According to the qualitative literature, French competition increased in the 1740s and the 1750s until it was suspended in the late 1750s by war. Based on the historical accounts and applying our methodology, we have derived a population series and a pattern of French harvest consistent with the observed trade at York Factory. Our simulations highlight the impact of French competition on depletion. The trade at York Factory was consistent with maximum sustained-yield management until 1739 when we estimate the beaver population at 116 thousand. Starting in 1739, however, the population fell sharply. The decline was due in part to exploitation by the French but even more important was the apparent change in strategy by the Hudson's Bay Company in response to the competition. After receiving about 30 thousand furs per year over the period 1732-8, which stabilized the beaver population, York Factory took in an average of 38 thousand furs over the years 1739-42,

reducing the beaver population to 75 thousand by 1743, and substantially lowered future returns from the region.

The third post, Fort Churchill, was furthest along the Hudson Bay coast and the least open to competition from the French. Here, we estimate the beaver population to have been much more stable through the period examined. In fact, prior to 1750, the population never fell below 80 percent of the level consistent with maximum sustained yield. This is very similar to the York Factory experience before it faced French competition in the 1740s. The trade at Fort Churchill provides further evidence that French competition was a major factor in the over-exploitation of the beaver population. As long as the Hudson's Bay Company had a monopoly, the harvest was at a rate approximately consistent with a maximum sustained yield. This was true at York Factory and Fort Churchill. It was only when the French began trading in a region that excessive harvests were observed. Also significant is the fact that these excessive harvests were due not to large numbers of furs being taken by the French, but rather to a shift in Hudson's Bay Company policy. In response to the French trade, the Company began to receive furs at much higher rates, rates that led to severe depletion of the beaver stocks. Only at Fort Churchill, which was insulated by distance from significant competition did Company policy not change.

In the above analysis, we assume that the Hudson's Bay Company could control the number of furs traded at each post, but given the nature of its relationship with the Indian trappers and middlemen, it could do so only indirectly. After all it was the Indians who did the actual trapping. The Hudson's Bay Company, however, could set fur prices through its post managers and it was pricing policy that ultimately determined, or at least strongly affected, the size of the fur harvest. To derive a relationship between fur prices and the harvest, we relate the (Indians') level of harvesting effort to the price of furs and the size of the resource stock. This allows us to relate price changes at the three trading posts in our study to changes in our estimated beaver population. Throughout the period 1701-63, although the Company's Official Standards, which remained constant over time and was the same across posts, actual prices received rarely equalled that standard because post managers were permitted to offer the Indians lower prices, and record an overplus in the accounts. At the same time, managers incurred "expenses" in addition to the value of goods traded directly for furs. These expenses represented the cost of gifts given to the Indians, and were an important feature of the trade. When the overplus exceeded expenses, Indians received less than the official standard for their furs, and when the reverse was true they received more. Since both the overplus and the expenses are reported in the annual accounts for each post, it is possible to derive price series for each area.

A comparison of the Fort Albany and York Factory price series with the population estimates suggests a connection between the prices Indians received for their furs and the rate of depletion. In the Fort Albany region we estimate that from 1720 to 1730, when the beaver population was stable, there was almost no change in fur prices; whereas during the 1730s, when fur prices were raised by about 20 percent, the beaver population declined from 60 to 50 thousand, or by about 15 percent. There was another large price increase starting in the late 1740s which saw the price of furs move from .9 to 1.25 MB by 1755, and the beaver population fell by more than 50 percent. This negative relationship is also observed in the York Factory data. In that region the population fell from 115 thousand in the late 1730s to 50 thousand by 1750, with fur prices increasing from .68 to .99 MB. Equally suggestive is the evidence from Fort Churchill, where we estimate the beaver population remained fairly stable throughout the period, and the price series for Fort Churchill is consistent with that stability.

Our analysis highlights the role played by competition in the over-exploitation of the beaver population. As we point out, what was important was not the size of the French harvest but rather the fact of the French presence and the Company's reaction to that presence

SESSION THREE - CONTRACTS AND INSTITUTIONS
SATURDAY, 2:30 P.M.

THE MERCHANT GILD AS A NEXUS OF CONTRACTS

Avner Greif Paul Milgrom Barry Weingast
Stanford University

The process of European economic growth from the eleventh to the fourteenth centuries is best described using the words of the great economic historian Robert Lopen as "The Commercial Revolution of the Middle Ages." Economic growth was based upon commercial expansion. Many studies have outlined the nature of this commercial expansion and the associated development of trade practices. For this commercial expansion to be possible, however, institutions capable of mitigating contractual problems associated with long-distance trade were required. The operation of these institutions was more important than might be inferred from a simple reading of the historical record. The effectiveness of institutions for deterring breach of contract might best be judged like that of peacetime armies - by how little they are used. In reading the historical record to determine whether institutions like the merchant gild were created largely to ensure contract compliance, the number of instances of enforcement is not a useful measure. Instead, one has to ask: What were the things that threatened, and on occasion thwarted, efficient trading? Can details of these institutions be explained as a response to those threats? What triggered major changes in these institutions? Were changes in structure followed by important changes in trading relations, especially by improved contract compliance?

A comprehensive analysis of an enforcement institution must consider why it was needed, what sanctions were to be used to deter undesirable behavior, who was to apply the sanctions, how the sanctioners learned or decided what sanctions to apply, why they did not shirk from their duty, and why the offender did not flee to avoid sanction. Some full game-theoretical analyses according to these criteria have been developed. Greif (1989) has analyzed the institution that governed the relations between merchants and their overseas agents in the eleventh century Mediterranean trade. To reap the benefit of employing overseas agents, a specific group of merchants, known as the Maghribi traders, organized agency relations with a "coalition" whose members ostracized and retaliated against agents who violated their commercial code. Interrelated contractual arrangements assured proper incentives, while close community ties assured that each member had the necessary information to participate in sanctions when necessary. Milgrom, North and Weingast (1990) have argued that the use of merchant courts in the Champagne fairs can be analyzed as an institution with proper incentives for gathering information, honoring agreements, reporting disputes, and adhering to the judgments of the merchant courts. Moreover, by centralizing certain record keeping functions and effectively permitting only merchants in good standing to remain at the fairs, this institution also achieved significant transaction costs economies relative to other feasible enforcement institutions.

The papers we have just cited provide consistent analyses of institutions used to overcome contractual problems between individual merchants active in long-distance trade. The merchants, however, were not the only important parties; the trading centers where the merchants met and brought their goods were an important independent force. This paper is concerned with the institutional arrangements that surmounted the commitment problem associated with the relations between a ruler and alien merchants. It develops a theory to explain how the attributes of the institutions that mitigate the ruler's commitment problem

affect the possibility of trade expansion. This theory is applied in a preliminary historical study of the medieval Merchant Gild.

In his own territory, the Medieval ruler had the ability to abuse the rights of alien merchants using his coercive power. In the age before the emergence of the "state," alien merchants could expect little military or political aid from their countrymen. Without any institutional arrangements enabling a ruler to commit himself ex-ante, not to abuse alien merchants' rights ex-post, after the merchants had come to trade, the merchants were not likely to frequent that ruler's territory. Thus, trade could not expand, leading to the deprivation of rulers and traders alike.

Intuitively, one can conjecture that the rulers' commitment problem can be surmounted by bilateral or multilateral reputation mechanisms in which the threat of the merchants to impose an embargo on a trade center that abused their rights is enough to deter abuse. However, when the ruler-merchants commitment problem is modeled, this intuition is found to be inaccurate. Simple bilateral and multilateral reputation mechanisms, without supporting institutions, lack the attributes required to overcome the commitment problem at the efficient level of trade. Accordingly, the model is used to study the attributes required from institutions to support an efficient level of trade by supplementing the operation of multilateral reputation mechanism. The multilateral reputation mechanism in which the merchants are to collectively punish a ruler who abused the rights of some of them to surmount the ruler-merchant commitment problem must be a part of an institution that has the ability to coordinate merchants' responses and to force all (or most of) the merchants to follow its decisions after the rights of some merchants were abused. An institution with these attributes may enable trade to expand to its efficient level by surmounting the commitment problems.

Bilateral and multilateral reputation mechanisms fail to surmount the commitment problem without supporting institutions since in the efficient level of trade the "value" to the ruler of the "marginal traders' trade" is "very low." Thus the threat of future boycott by these marginal traders would not prevent the ruler from abusing their rights. In other words, when the ruler-merchants relations are governed by bilateral or multilateral reputation mechanisms there is no sub-game perfect equilibrium in which the ruler's "promise" to respect the rights of the merchants in the efficient level of trade is credible and hence trade can not expand to its efficient level. On the other hand when a multilateral reputation mechanism is supplemented by a "coordinating" mechanism, that is, a mechanism that coordinates the responses of all the merchants to the abuse of the rights of some merchants, the commitment problem can be overcome. There is a sub-game perfect equilibrium in which the ruler's "promise" to respect the rights of the merchants is credible in the efficient level of trade.

When a coordinating mechanism exists there is an equilibrium in which the commitment problem is surmounted. It is the efficient equilibrium but not the only equilibrium possible. Accordingly, we use the notion of re-negotiation proof to evaluate the plausibility of this equilibrium and ascertain that it is not a plausible equilibrium. Intuitively, the coordinating institution's threat to punish a ruler by declaring an embargo if he ever abuses rights is not credible since some merchants will find it to their advantage not to respect the coordinating institution's embargo decision. If the number of these merchants is large enough, the embargo threat is not credible and the commitment problem can not be surmounted through a coordinating institution. Thus, to overcome the commitment problem there is a need for an institution that also possesses "internal enforcement" attributes. That

is, an institution that has the ability to force its decisions regarding their relations with rulers upon the merchants.

The historical analysis of the commitment problem led us to advance the hypothesis that during the late Medieval period a specific institution, the Merchant Guild, functioned as a nexus of contracts that surmounted the ruler-merchants commitment problem by supplanting the operation of multilateral reputation mechanism with its ability to coordinate and enforce merchants' responses. The core of a Merchant Guild was a society of merchants organized on territorial bases with some regulatory power over its members. This regulatory power over its members. This regulatory power could not be subverted by the ruler, usually because the Guild's territorial bases were out of his reach. The Guild governed the relations between an individual merchant and rulers of regions outside the Guild's territorial bases, and the relations between the merchants themselves. The Guild established procedures for information gathering and transmission that coordinated the responses of all the merchants in cases when the rights of some of them were abused abroad. The guild used its power to regulate merchants' activities in its territorial bases to insure its members' obedience to the Guild's decisions. In particular, it insured that all merchants observe trade embargoes declared by the Guild against foreign rulers who had abused members' rights. The Merchant Guild appeared during the late Medieval period in various forms -- the German Hansa, the English Guilds and some Italian and Germanic towns -- were Merchant Guilds. They were merchant guilds because they provided the merchants with the mechanism required for coordination and internal enforcement.

By fulfilling these functions of coordination and internal enforcement, the Guild enabled the merchants to collectively commit themselves to retaliate by never trading again in a city whose ruler had abused the rights of any individual member. This collective credible threat increased the long-run cost of abusing traders' rights above the short run gains from abusing. Thus, it became known ex-ante that the best a ruler can do ex-post, after the alien merchants had arrived in his city, was to respect the merchants' rights. Thus, the merchants could "trust" the ruler ex-ante and trade could expand. The Guild was an institution in the sense that its operation constrained the ruler and the merchants by altering the payoffs associated with specific actions. By channeling their choice of behavior, the Guild enabled them to commit themselves.

To support the above hypothesis about the role of the Merchant Guilds in overcoming the commitment problem this paper provides historical evidence. Since, in other times and places, the details of Guilds' organizations differed, we distinguish between three groups of historical evidence in our presentation.

The first group contains evidence on the importance of the commitment problem, the role of reputation mechanism in overcoming it, and the relations between overcoming the commitment problem and trade expansion. The relation between the credibility of the commitment to respect alien merchants' rights and trade expansion seems to have been clear to Medieval rulers. For example, Edward the Third, King of England, declared in 1283 that because alien merchants did not receive the protection they had expected, "many merchants are put off from coming to this land with their merchandise to the detriment of merchants and of the whole kingdom." Edward's words should be understood against the background of events like the one that occurred in Boston, England. According to an enquiry conducted in 1241, after a Flemish merchant accused an English trader of not repaying a commercial loan "there was an uproar on all sides and the English merchants assembled to attack the Flemings, who retired to their lodging in the churchyard,...The English threw down the pailings, broke the doors and windows and dragged out Peter Balg [the lender] and five others, whom they foully beat and wounded and set in the stocks. All the other Flemings they beat, ill-treated and robbed, and pierced their cloths with swords and knives. Their silver cups were carried off as they sat at table, their purses cut and the money in them stolen,

their chests broken open and money and goods, to an unknown extent, taken away, and their cloths pierced with knives."

The historical records suggest that the evolution of guilds facilitated the growth of trade. For example, it was noted that the Catalan merchants' trade expanded "within only a few months" after they received, in 1286, privileges and the right to have a consul in Sicily. The trade of the German merchants in Bruges expanded after they had received privileges and the right to have a guild, called a *Kontor* (establishment or office). The fact that the Italians were able to secure their trade in Flanders only after they established guilds, called 'nations,' indicated the importance of these organizations in increasing the ability of the merchants to retaliate.

The importance of reputation in providing protection to alien merchants is reflected, for example, in the agreement made in 1261 between Flemish merchants from different towns who were "associated with the buying of wool from abbeys" in England. The agreement reflects their decision that "if it should happen that any cleric or any other merchant anywhere in England who deals with sales of wool deals falsely with any merchant in the alliance, by giving false weight or false dressing of the wool or false produce,...and if they do not wish to make amends, we have decided that no present or future member of this alliance will be so bold as to trade with them..." Note the implied inability to seek justice in the local court and the use of collective economic sanctions to increase the cost of cheating any individual merchants.

The second group of evidence contains information about the internal structure of Guilds, the contractual arrangements between Guilds and their members, between Guilds and the rulers of the Guilds' territorial bases, and between Guilds and rulers who controlled trade centers outside the Guilds' territorial bases. The evidence indicates that the attributes of the Guild and its contractual arrangements were parallel to those found in our theoretical analysis to be required to overcome the commitment problem. For example, the Flemish merchants whose agreement was quoted above recognized that some coordination is required to make the boycott functional and hence they "have decided that there will be in each of these cities one man to view and judge the grievances, and to persuade the wrongdoers to make amends."

Simple coordination, however, was not usually sufficient and internal enforcement was required. Historical evidence indicates that ignoring a ban could be very profitable and that guilds faced difficulties in insuring that trade embargoes would be respected by their members. In 1284 a German trading ship was attached and pillaged by the Norwegians. The German towns responded by imposing an embargo on Norway. The export of grain, flour, vegetables and beer was prohibited. According to the chronicler Detmar, "there broke out a famine so great that (the Norwegians) were forced to make atonement." The temptation for an individual merchant to smuggle food to Norway in this situation is clear. To prevent smuggling and sustain the embargo, the German towns had to post ships in the Danish Straits. The fact that the success of a trade embargo depended crucially on obtaining the support of virtually all of the individual merchants involved was also clear to the cities on which embargo was inflicted. When, in 1358, the German towns imposed an embargo on Bruges, the city attempted to defeat the embargo by offering Cologne extensive trade privileges.

Similar attention to the need to guarantee solidarity of incentives among merchants is also reflected in Flemish regulations for trade at the English fairs written in 1240. "If any man of Ypres or Daouai shall go against those decisions [made by the guild]...for the common good, regarding fines or anything else, that man shall be excluded from selling, lodging, eating, or depositing his wool or cloth in ships with the rest of the merchants...And if anyone violates this ostracism, he shall be fined 5s..." Placing ships in a strait and imposing fines are specific ways to overcome the distinct incentives problem.

The evidence, however, indicates a variety of means through which the credibility of the threat to carry out an embargo was sustained.

In England, for example, a gild had exclusive trade privileges in its own town, usually including monopoly rights over retail trade within the town, exclusive exemption from tolls and so forth, as well as the right to exclude under certain circumstances members from the gild. The English gilds were thus able to provide their members with streams of rents in their home towns. Receiving these rents, however, was conditional to following the recommendations of the gild as the Flemish regulations of 1240 illustrate. A merchant who ignored the ban imposed by the gild on another town was expelled, losing his rent stream. Hence, the Gilds' exclusive rights should be understood as an integral part of a system that facilitated trade expansion rather than being simply monopoly rights that hindered trade.

Finally, the paper provides a description of the evolution of a specific Gild - the German Hansa - during the thirteenth and fourteenth centuries. This description illustrates that the evolution of this Gild is consistent with our theoretical analysis -- to overcome the commitment problem, for the ruler to commit himself to respect the merchants' rights, there is a need for an institution with coordination and internal enforcement abilities that control a sufficiently large number of merchants. For historical reasons, in the basic organizational unit that coordinated the activities of German merchants abroad -- the *Kontor* -- membership was not conditional upon residency in one particular town. Any German merchant who arrived in a non-German city could join the local *Kontor*. While the *Kontor* had the same function as the gild in coordinating the responses of the German merchants in disputes with the town, it lacked the enforcement ability against its own members. This weakness, and the way it was overcome, is the essence of the history of the contractual relations between the German *Kontor* of Bruges, and the German towns.

In 1252, a *Kontor* of the German merchants obtained extensive trading privileges from Bruges, and a permanent settlement followed. Before long, however, the trading privileges given to the alien merchants in Bruges were continually abused and eventually the German merchants under the leadership of Lubeck retaliated in 1280 by transferring their trade from Bruges to Aardenburg. After two years of negotiation, a new agreement was reached and the alien merchants returned to Bruges. Only seemingly successful, the embargo failed to guarantee the property rights of the German merchants in Bruges, as the city simply ignored its agreement with them. It should be noted, however, that Bruges did respect the rights of other alien merchants who frequented the city. Our analysis points to the reason for that uneven treatment. The embargo was not imposed by the German merchants but by the alien merchants in Bruges in general. In particular, the important and well organized Italian and Spanish "nations" took part in the boycott. While the lesson for Bruges from that episode was to respect the rights of those well organized nations, it soon became clear to the city that the German merchants were the "marginal" traders. The *Kontor* was not capable of increasing the cost of abusing German merchants since it did not have the ability to enforce its own decision upon its members. A main source of this inability was that the *KONTOR* embodied only the German merchants actually present in Bruges rather than all the potential German traders.

Another embargo, from 1307 to 1309, in which only the Germans participated was required to force Bruges to respect its contractual commitments with them. Between 1280 and 1307 the ability of the German traders to coordinate their responses on the level of towns was enhanced as a consequence of the embargo they imposed in 1284 on Norway. The towns collectively achieved for the first time the coordination needed to expel one of their members. Hence, by 1307, the ability of the German merchants to commit themselves, to coordinate their actions, and thus guarantee Bruges's adherence to its contractual obligations, was rather advanced although it was still the local *Kontor* that organized the

embargo. Bruges respected the charters agreed upon in 1307 and 1309 and, consequently, Flanders' trade flourished and expanded for the next 50 years.

It was not until the middle of the century, when the cost of providing security around Bruges rose drastically that a new level of cooperation among the German towns was needed to force Bruges to provide the security required to support trade. The Hansa relations with Bruges deteriorated around 1350, mainly because Bruges was not ready to compensate the Germans for their damages in Flanders from the war between England and France. The Hansa responded by strengthening its internal organization and held, in 1356, its first *Diet* in which it was agreed that the *Kontor* of Bruges should be operated according to the decisions of the *Diet*. The institution of the German Hansa was not crystallized and an Hanseatic trade ban followed in 1358. It was announced that any disobedience, whether by a town or an individual, was to be punished by perpetual exclusion from the Hansa. This time, recognizing the ability of Hanseatic towns to coordinate and internally enforce their decisions, the privileges were written, "in much detail as to prevent any one-sided interpretations." The trade of Northern Europe prospered under the supremacy of the Hansa, and although the trade embargo of 1307 was not the last, later trade disputes seem to have been centered around distributive issues - the provision of trade privileges. Commitment for security was no longer an issue.

It is illuminating to compare the development of the Hansa among German towns with the rather different organization among the Italian cities. The solid internal, political and commercial organization of the Italian cities and their prominence in trade enabled them to overcome the coordination and distinct incentive problems. None of the cities were "marginal players" in the ports where they traded. In contrast, the German *Kontor* was a local organization in a trading center which lacked the ability to enforce its decisions upon its members who came from various German towns. The German towns themselves were small and, before the establishment of the Hansa, each of them was relatively insignificant in its trading relations. For these reasons, overcoming the commitment problem required a trans-town organization. As described above, these towns eventually formed the Hansa in order to present a more unified front against the cities where they traded. The resulting increased security afforded to the merchants and their property contributed to increasing levels of trade in the regions.

The historical application of our theoretical model is to the historical issue that led to the construction of this model, namely the study of ruler-merchants relations. We contend, however, that the insights of this model can be used to highlight the essence of other institutions that throughout history advanced efficiency by enabling rulers to commit themselves.

GOVERNMENTS-AS-FIRMS: MORE ON THE EVOLUTION OF PATRONAGE

Joseph D. Reid, Jr. Michael M. Kurth
George Mason McNeese State
University University

Governments increasingly attract the attention of economic historians. Unfortunately, often governments are mistaken for government. The implied assumption being that there is a political sphere and there is a nonpolitical sphere, and in charge of the political sphere is government. In history, however, neither the sphere of governments nor their organization are well settled. In one sense, we all know this: after all, the economic history of the United

States cannot antedate the United States, and so begins at earliest with revolution in 1776. But in another sense we don't know this: economic historians have been slow to grasp that government is what governors make it, and in the large governors have tried steadily to expand their government, at the expense of the private sector and of other governments.

In American history, horizontal and vertical competitions among governments are well mirrored by the evolution of organization of employees of federal, state, and local governments. In each government, employees were organized into patronage and then into civil service bureaus. But at each level these evolutionary paths seem dissimilar and diverse--urban patronage grew while federal patronage declined, for instance, urban patronage mainly was in industrial rather than in agricultural areas, and later state and local unionism grew while federal unionism stagnated--but in fact all of these changes are remarkably well described as outcomes from attempts by political firms or would-be governments to minimize cost and maximize payoff from governing, while increasing the costs and reducing the payoffs to competitors. It seems that government employees are employed by politicians to best produce current and future income and votes, and politicians are the cooperating entrepreneurs in political firms designed and managed for maximum expected political present value through entrenchment and expansion.

I. Theory

As with any firm, public or private, would-be or established, maximization of expected present value requires choice of what to produce, how to produce, and how to sell. Because pluralities of votes capture political markets in the United States, the first task of an American political firm is to find out what is wanted by alternative minimum winning coalitions of voters, and to supply the most profitable coalition. This is, admittedly, no easy task, since voters' interests rationally are masked by their reluctance to spend effort or income for publicly provided goods. Furthermore, since generally there are lead times to supply voters' wants, the evolution of minimum winning coalitions should be anticipated and, if possible, affected by policy. In sum, what is best for political firms to sell requires discernment of what voters will pay for from what voters say they want, and requires forecast and market direction. This means that at any time there will be confusion and there will be failing political firms. It further means that when voters' wants are clear and stable, most political firms will promise the same things and will compete on costs. But when voters' wants are murky or unstable, political firms will promise different things, including to lead (in anticipation of growing consensus) or to follow closely and efficiently (in hope of staving off challenge from newly emerging coalitions). When different governments face different costs or constituencies, they understandably will supply different things. And if one government can expand with profit at the expense of another, it will.

There are entrepreneurs and hack producers in public as well as in private firms, in short. But there are differences. Choices of how to produce and sell are more important to political firms than to private firms, because quids are more separated in time from quos and are more varied in political exchanges. A typical private market sale is the customer gets the good at the same time the provider gets the sale price in legal tender. A typical political market sale is the customer gets the good and much later the government gets paid in varied currency (votes, in-kind campaign assistance, legal tender). If a government bought a promised lifetime of votes by creating a long-lived source of voter payoffs (a dam, for instance), voters would withhold their votes without fear of loss of the benefits (that flow for a long time automatically); thus "what has politician X done for me lately?" is a stereotypical illustration of voters' fickleness. Other things equal, then in political markets sales will be small and continuous to voters, in order to link payment to payoff. On the other hand, if a government sold a promise of production of a special interest's want, it could sell its promise for more the more irreversible the outcome; thus a constitutional amendment would

sell for more than a statute, and a rule from an independent agency would sell for more than a promise of favor upon request, no matter how fervently made, for who knows how circumstances will change, but circumstances will little affect outputs from independent agencies, the Constitution, and large capital intense (low marginal cost) sources (such as dams).

Technology as well as honor of parties helps to determine how government transactions are produced and sold. If technology increasingly favored capital intensive means of production, then competition on costs would prompt more capital intensive production, other things equal. If particular voters could be better identified, so that their demands and payoffs could be measured more accurately, then these voters would gain more payoff from politicians. These insights are illustrated by the evolution of governments in the United States. By removing the South as an effective blocker of federal growth, The Civil War enlarged the potential domain of the federal government. The circumstances were most propitious, for demand for that most American of government activities, subsidization of transportation improvements, had been pent up by Southern reluctance and then by absorption in the War effort. The lands that railroads would make valuable were homogeneous and unpopulated, and were championed by well organized special interests, as well--railroads and coastal merchants anticipating profits from trade from the interior. Homogeneity meant that large-scale solutions to preparing the interior for settlement would be economical, and lack of western settlement meant that voters' payoff for improvements need not be linked to delivery; a few easily monitored railroads would pay. Thus, seeking economies of scale and unconcerned about discovering and collecting the value of western improvements from settlers, the federal government moved to produce political outcomes independently of payment or response of constituents. One means was through bureaucratization of independent agencies dedicated to production only.

At the same time, cities handy to bulk inputs (say, on water), run-down housing (old, on water) to house poor laborers, and fast output dispersion (on rail lines) began steam-powered industrialization and polyglot population growth. Municipal governments there faced demands from well placed special interests for commercial favors (to permit more smoke and more noise from factories, and to crowd adequate labor forces nearby) and demands from new potentially minimum winning coalitions for helps in coping with the risks of industrial work in a strange land--demands for help in finding lodging, jobs, and insurance from familial inattention, injury, or death. The demanded commercial favors were idiosyncratic and expectedly would be opposed by rivals, as well as by old line residents who fondly recalled quieter times. The demanded voter favors were equally idiosyncratic, in that relief for a large family with a badly injured breadwinner would be larger and longer than relief for a widower, and were subject to malfeasant overindulgence unless monitored. Thus, newly winning governments in such cities became patronage governments that so-to-speak sold idiosyncratic zoning variances to manufacturers, and delivered and monitored social insurance to poor workers in return for their votes.

In contrast to postbellum manufacturing cities, the mercantile wants of farm cities were homogeneously focused on efficient service of the crops. Where there was a dominant market crop, as in prairie agriculture, this was truest, because everybody was benefitted by efficient service of the crop. So, while increasingly diverse manufacturing cities turned to the neighborhood politics of patronage, homogeneous farm cities followed the federal government into the apolitical bureaucratized and professionalized politics of commission governments.

II. History

Technology and history raised the relative cost of patronage to reform government. Technology drove down the cost of communicating and tracking people and groups, while

passed time and lessened immigration homogenized and enriched urban populations. Urban homogenization propelled the replacement of patronage workers by bureaucrats in manufacturing towns, and the narrowing of membership in urban minimum winning coalitions to the great median of the white, stable middle-class

The Depression and the Second World War legitimated an expanded federal government. But the War's end left it with little legitimate to do. More likely by accident than design, the federal government stumbled into helping the downtrodden excluded from the new urban consensus. First helped were blacks. Blacks already had a means of organizing and articulating their interests, the black Christian churches. But technological improvements (such as computerized mailing lists) facilitated cheaper federal and state communication with other demographic or zip-code groups excluded from winning urban political coalitions. Reaching out for new constituents, the federal government aided those excluded from winning urban coalitions.

To spend on Blacks and others excluded from urban ruling coalitions secured the federal government a taxes shelter and a breach in limits on federal authority that was welcome. Many federal spending programs were institutionalized into categorical aid (revenue sharing) programs that directed and subsidized municipal spending. The trillion dollar subsidization of revenue sharing (between 1960 and 1980) was welcome. But the direction was unwelcome to the municipal governments that, had they wanted to embrace the excluded, earlier could have done so. To protect their turf and to satisfy their electoral coalitions, urban governments began to tolerate militancy among their public employees. Suddenly, job actions and work slow downs by municipal employees dispensing categorical shared revenues led

REGULATION AS A CONTRACT: THE ORIGINS OF THE ILLINOIS PUBLIC UTILITIES COMMISSION

Werner Troesken
George Mason University

State public utility regulation was never an historical necessity. Throughout the nineteenth century only one state, Massachusetts, regulated public utilities (excluding railroads) through a state commission. However, by the second decade of the twentieth century, this situation began to change. In the fifteen years between 1907 and 1922 more than forty states created public utility commissions.¹ What was happening to cause such profound institutional change? This paper provides at least a partial answer. Interpreting state regulation as a contract, it explains the emergence of the Illinois Public Utilities Commission in 1913. The primary source of evidence for this study is the late nineteenth and early twentieth century Chicago (manufactured coal) gas industry.²

Asset Specificity and Credible Commitments

The nature of gas production and distribution made it necessary for producers and consumers to invest in specific assets. Gas production required a large investment in fixed plant and capital which was not easily resold, moved or adapted. Consumers also needed to purchase unique and immobile assets such as gas lighting or heating fixtures. Beyond this, imperfect housing and labor markets further restricted consumer mobility. This asset specificity made it necessary for consumers and producers to enter into long-term binding rate agreements. Before they would make the requisite investments, producers needed to be

confident that they could set prices which would afford them sufficient rates of return. Likewise, consumers would not acquire the necessary fixtures before they were certain that producers were not going to charge monopolistic rates or provide inconsistent service. Public utility companies and consumers, in other words, must have been able to make credible commitments.³

Nineteenth Century Institutions

During the nineteenth century, gas producers and consumers in Chicago made credible commitments through municipal contract ordinances. In municipal contracting, gas companies bargained directly with the city council, which represented the interests of consumers.⁴ It is important to stress that in these proceedings the city did not unilaterally determine the terms of the contract; producers had to agree to the contract ordinance, in full, before it became binding. Through municipal contract ordinances the city granted gas companies certain operating rights. In return for these rights, producers agreed to a rate schedule setting consumer gas prices for the next five years. Typically, producers also agreed to pay the city a percentage of their earnings. Presumably this was done to compensate the city for the costs which resulted from gas companies digging-up the streets.⁵

For most of the nineteenth century three institutions- the market mechanism, Illinois law and transaction costs - promoted the efficacy of municipal contracting. Until 1897 competitive forces protected consumers against producers' attempts to monopolize the industry. Between 1880 and 1897 ten new companies entered the Chicago gas industry. Real gas prices fell by over 50 percent. In short, the market mechanism forced producers to agree to, and, in fact, credibly commit to, competitive rates. Simultaneously, Illinois' legal environment and high transaction costs constrained the behavior of consumers, forcing them to commit to gas prices which allowed producers reasonable rates of return. Transaction costs prevented consumers from organizing a monopsony and seizing the surplus of producers. Also, Illinois law severely restricted the regulatory powers of municipal authorities. The state and federal courts made it clear that without an enabling act from the Illinois legislature, the Chicago city council had

no power to unilaterally dictate gas rates.⁶ In future years, when this constitutional constraint was removed, voting consumers mandated that the city council pass confiscatory rate ordinances. The effectiveness of these three institutions limited the demand for state regulation. As a consequence, all attempts to create a state public utility commission prior to 1900 failed.

Legal Change and the Demise of Municipal Contracting

At the turn of the twentieth century, three legal changes combined to destroy the institutions forestalling opportunistic behavior. First, around 1900, in *Rogers Park v. Fergus*, and several other decisions, state and federal courts began ruling that municipal contract ordinances were unenforceable.⁷ As a result of these rulings, Kneier explained that Illinois "cities had no power to make rate contracts for a long period of years."⁸ So even if producers and the city could agree on rates *ex ante*, enforcing such agreements through the courts would have been difficult.

The second institutional change hastening the demise of municipal contracting was the passage of the Lowenthal street frontage act and the gas consolidation act in 1897. The consolidation act removed the common law obstacles preventing Chicago gas companies from merging into a single firm. The Lowenthal frontage act erected an impenetrable entry barrier.⁹ After 1897, not a single new firm successfully entered the industry. Gas prices rose

substantially. By 1913, the Peoples' Gas Light and Coke Company was the sole supplier of gas to the city. In effect, the gas acts created and sanctioned monopoly, destroying the one institution - the market mechanism - which prevented producers from charging monopolistic rates.

Finally, in 1905, Illinois passed legislation empowering the Chicago city council to regulate gas rates. The law no longer provided gas companies clear protection against municipal rate regulation.¹⁰ After 1905 consumers possessed the political power to influence prices. With a mandate from voters, the city council would pass coercive and confiscatory rate ordinances.

The gas acts of 1897 and the enabling act of 1905 greatly expanded the rate-setting powers of both consumers and producers. A Bilateral monopoly situation emerged, replacing one where neither party and enough economic or political power to substantially influence rates. As a result, it became far more difficult for producers and the city council to agree on rate schedules *ex ante*. Moreover, even if they could design such an agreement, court enforcement would have been problematic in light of *Rogers Park v. Fergus* and other similar decisions. By 1911, bargaining between gas producers and the Chicago city council collapsed.¹¹ Untimely and costly litigation followed.¹² State regulation quickly became a more compelling alternative.

The Origins and Efficacy of State Regulation

In an effort to replace the institutions which prevented opportunistic behavior for most of the nineteenth century, Illinois created the Illinois Public Utilities Commission (IPUC) in 1913.¹³ State regulation acted as a self-enforcing contract, enabling producers and consumers to credibly commit. In the absence of competitive forces, the IPUC protected consumers against attempts by gas companies to charge monopolistic rates. Similarly, as constitutional constraints had done prior to 1905, state regulation prevented consumers, acting through the city council, from taking the property of producers.

The legislative history of the Illinois Public Utilities Act, the act which created the IPUC, supports this contract interpretation. Support for state regulation among utilities, though no ubiquitous, was strong.¹⁴ Moreover, producers typically demanded state regulation because they believed it would protect them against the relatively hostile regulation of municipalities.¹⁵ Chicago consumers, in contrast, generally desired a regulatory framework which vested the city council with supreme regulatory powers, not a state-wide commission. That is, consumers wanted "home-rule."¹⁶ Chicago utilities correctly feared such action and were "elated" when the legislature deleted the portion of the Public Utilities Act granting Chicago home-rule.¹⁷ The actual performance of the IPUC further suggests that state regulation was designed to protect both consumers and utilities. Gas prices in Chicago, and elsewhere in Illinois, were not materially altered, in either direction, after the creation of the IPUC in 1913. Moreover, both accounting data and market data suggest that Chicago gas producers did not reap large profits under state regulation.¹⁸

problems which pervaded the gas industry were also present in other public utility industries. For information on the proportion of public utility consumers residing in Chicago at the turn of the century, see Majority and Minority Report of the Special Committee on Public Utilities of the Forty-Ninth General Assembly of the State of Illinois, January 20, 1917, p. 20. Hereafter this citation will be referred to as Majority and Minority Report.

³This argument was initially advanced by Goldberg. See Victor Goldberg, Regulation and Administered Contracts, 7 Bell J. of Econ. 426 (1976). See also, Oliver Williamson, The Economic Institutions of Capitalism (1985).

⁴Standard historical sources suggest that the late nineteenth century Chicago city council was corrupt. Nonetheless, the history of the Chicago gas industry indicates that the city council was generally more sympathetic to the interests of consumers than to the interests of public utilities. The full-length version of the paper examines this issue in more detail. I thank Louis Cain for his comments and for directing me to several helpful sources on Chicago history.

⁵For descriptions of these contracts see Comm. & Fin. Chron., June 20, 1891, p.939 and Comm. & Fin. Chron., May 25 1895, p. 928.

⁶For example, in 1900 the Chicago city council passed a coercive rate ordinance, *not* a contract ordinance, requiring gas companies to reduce their rates from \$1.00 per 1,000 cubic feet to \$0.80. Darius Mills, a stockholder of one of the larger Chicago gas companies, sued for injunctive relief. In *Mills v. Chicago*, a federal court ruled; "the regulation of prices to charge consumers by gas companies is not one of the powers essential to municipal government, and is not included in general powers conferred on cities..." The court continued: "and such power cannot be exercised by a city unless it has been delegated by the state in express words..." The ordinance was not enforced. See *Mills v. City of Chicago et. al.* 127 Fed. 731 (1904), p. 731.

⁷In *Rogers Park v. Fergus*, the Illinois Supreme Court ruled: "An ordinance granting a corporation the right to use public streets for that period is merely a declaration...and is not a contract which binds the city to recognize such rates...for the full period." See *Rogers Park Water Co. v. Fergus*, 178 Ill. 571 (1899), p.571; *Affid. Rogers Park Water Co. v. Fergus*, 48 L. Ed. 702 (1901). See also Herbert Pope, *Municipal Contracts and the Regulation of Rates*, 16 Harv. L. Rev. 1 (1902) for a comprehensive survey.

⁸Charles M. Kneier, *State Regulation of Public Utilities in Illinois*, 14 University of Illinois Studies in the Social Sciences 9 (1926), p. 68

⁹The Lowenthal street frontage act, which was originally introduced by a state senator from Cook County (Chicago is located in Cook), provided that no city or town in Illinois could grant a franchise "for laying of gas pipes...without the consent of the owners of more than one-half of the property fronting the street or alley along which it [was] ...proposed to lay the pipes..." *Amer. Gas Lght. J.*, March 25, 1895, p.413. The frontage act also "made it possible for a person with the smallest interest in a piece of property facing a street or alley to go to court and block the laying of pipes..." Report of the Gas Bureau of the City of Chicago (1914), p. 22.

¹⁰Even after the passage of the enabling act in 1905, substantive due process provided public utilities some, albeit imperfect, protection against confiscatory rate regulation by either state or municipal authorities. See Herbert Hovenkamp, *The Political Economy of Substantive Due Process*, 40 Stan. L. Rev. 379 (1988) and N. Mathews, Jr. and W. G. Thompson, *Public Service Company Rates and the Fourteenth Amendment*, 15 Harv. L. Rev. 249 (1901). Undoubtedly, though, public utilities would have preferred regulatory environments which minimized their need to appeal to the courts for such protection.

¹¹In the spring of 1906, gas companies and the city council managed to agree on one last contract ordinance. They agreed to set rates at \$0.85 per 1,000 cubic feet (MCF) until 1911. In the spring of 1911, when the 1906 contract ordinance was set to expire, negotiations between the city and producers collapsed. For the first time the city council attempted to assert the regulatory power it received under the enabling act of 1905. The council passed a coercive ordinance requiring gas companies to charge \$0.75 (MCF). Producers stonewalled, refusing to lower their rates from \$0.85 (MCF). A complex tangle of battles wove through the state and federal courts. The dispute between the city and Chicago gas companies would not be fully resolved until 1927. See *Mills v. Peoples Gas Light and Coke Co.*, 327 Ill. 508 (1927).

¹ See 4 Public Opinion and Public Utilities, Speakers' Bulletin 1 (1920), esp. pp. 14-5 and George J. Stigler and Claire Friedland, "What Can Regulators Regulate: The Case of Electricity," 5 J. Law & Econ. 1 (1962).

²This case study approach does not severely limit the generality of the argument. Chicago has over one-half of Illinois' population. It had an even greater proportion of the state's public utility customers. For example, 70 percent of the state's gas consumers resided in Chicago. Furthermore, the contracting

¹² The Special Joint Committee, a committee created by the Illinois Legislature to study public utilities in Illinois prior to 1913, explained: "The effort of the city of Chicago as well as many other cities of the State is to regulate rates and services of utility companies by law suit. The purpose of litigation is merely to redress wrongs and not to afford a system of regulation. The method of thus supervising utilities is unscientific, expensive, vexatious and cumbersome... The termination of such litigation between municipalities and utility companies merely affords temporary relief... In Chicago there has been almost constant litigation between the city and utilities, resulting in heavy burdens of taxation being placed upon its citizens and an extraordinary expense being incurred by utility companies... The economic waste thus occasioned has been staggering. The present method of regulation as applied to the city of Chicago and the State of Illinois is disastrous alike to the public, to the utility companies and their patrons." Quoted from Report of the Special Joint Committee to Investigate Public Utilities, Journal of the Senate of the State of Illinois, April 17, 1913, p.860-1. Hereafter this citation will be referred to as simply Special Committee Report.

¹³ In 1913 the Illinois legislature wrote: "If municipalities are incapable of protecting their citizens for any reason from unjust exactions of public service corporations, it is the duty of the State to protect them...Conversely, if the citizens of any municipality, through their representatives, take such action as will destroy or confiscate public utility investments, it is likewise the duty of the State to assert its paramount authority to the end that justice may be accorded to citizens interested in such concerns..." Special Committee Report, p. 861.

¹⁴ See Forrest MacDonald, Samuel Insull and the Movement for State Utility Regulatory Commissions, 32 Bus. Hist. Rev. 241 (1968); Lloyd Wendt and Herman Kogan, Lords of the Levee (1943), pp. 172-3; Special Committee Report, p. 855; and Mansel Friffiths Blackford, Businessmen and the Regulation of Railroads and Public Utilities in California During the Progressive Era, 44 Bus. Hist. Rev. 307 (1970). See also the following issues of the Amer. Gas Lght. J.: Sept. 28, 1908, p. 537; March 25, 1912, pp. 207-8; March 27, 1911, P. 595; Oct. 12, 1908 pp. 620-1; and April 12, 1909, pp. 634-5.

¹⁵ In general, utilities were able to exercise more influence over state regulators than municipal regulators, because, city alderman represent smaller constituencies and act on a smaller range of issues than do state representatives. Voting consumers, in other words, held more power over municipal authorities than state representatives. Chicago alderman, and University of Chicago Professor, Charles Merriam explained: "The real reason why many corporations prefer state to local control is...that the indirect pressure of the state electorate is preferred to the direct pressure of the local electorate." Quoted from Majority and Minority Report, p. 27. Kogan and Wendt concurred when they wrote of late nineteenth and early twentieth century Chicago politics; "It has always been...strange...that a [state] legislator can be bought cheaper than an alderman." Kogan and Wendt (1943), p.172.

¹⁶ See Kneier (1926), p. 158.

¹⁷ During the final debates over the Illinois Public Utilities Act, the Springfield *Illinois State Register* reported that "it was quite significant during the fight that the corporation lobby vigorously opposed the 'home-rule' feature, and was elated when that principle was finally eliminated." Springfield Ill. State Reg., June 23 1913, p. 14.

¹⁸ See the full length version of the paper for a detailed presentation of this evidence.

"IF IT'S YELLOW, IT MUST BE BUTTER": MARGARINE REGULATION IN QUEBEC SINCE 1886

**Ruth Dupre
Ecole des Hautes Etudes Commerciales de Montreal**

In April 1918, the Canadian Minister of Agriculture, Thomas Crerar, told the House of Commons that "in the whole range of the business of agriculture, in all the countries where it is carried on, there is perhaps no question which has given rise to more controversy than the question of oleomargarine."¹ In Quebec, some seventy years later, it was still a matter for public debate when in August 1987 the government passed a bill stipulating that margarine

could not be yellow. Many people were puzzled but would not have been had they known the peculiar history of margarine regulation in the Western world.

The Poor's Butter and the Politician's Bread

Potato, bread, and soup were the staple food of the nineteenth century's poor in Europe. This diet may have provided sufficient carbohydrates but was quite unbalanced and particularly lacked protein and fat. The fat supply was not only insufficient, its composition was unsatisfactory as the meat fats-suet, lard, and bacon-, which made up close to 50% of total fat consumption were not easily digestible nor very nutritive. Besides, they could not be used as a spread. Butter was the answer: it could be spread and was relatively wholesome but its price was high, prohibitively high for most workers, and rising, by almost 100% between 1850 and 1870 in Europe.² There was thus much pressure to create a substitute for butter. At the 1866 Paris World Exhibition, Napoleon III offered to sponsor research to develop a cheap and nutritive fat. The French food chemist Mege-Mouries responded to the invitation and invented oleomargarine, made from beef tallow, in 1869.

This invention should have been most welcome and the margarine industry encourage. Yet, it is hard to find another industry which had to swim against such a tide of adverse government policy. Governments claimed they were protecting consumers from a product injurious to health or against its fraudulent sale as butter, but they had also the somewhat less noble objective of protecting dairy producers from a rival who could put on the market a substitute to butter at half its price.

The three objectives, often intertwined, did not require the same type of government intervention. If oleomargarine was really a poison deleterious to health, then nothing less than prohibition would do. This is what the extremists in the 1886 debates over the Canadian and American oleomargarine bills argued. They did everything to convince their colleagues, bringing chosen samples of margarine and reading disgusting descriptions of its fabrication using dead diseased cattle, horses, cats, and hogs floating into big vats. "It seems incredible now that intelligent people - legislators, in fact, should have believed [these] stories..." wrote Snodgrass (1930:37). Perhaps she was a little too optimistic. Some members of the Canadian government appeared to have taken these descriptions quite seriously.

Although the preamble to the 1886 Canadian Prohibition Act reads "whereas the use of certain substitutes for butter...is injurious to health", most participants in the debate did not consider oleomargarine was a dangerous product. It could be if improperly fabricated. In this respect, margarine was not any different from other food products like butter or meat which were not banned by merely regulated. By licensing, inspecting and setting up quality standards, the government could have supervised the manufacturing process while labelling and packaging norms could have taken care of the problem of fraudulent sale as butter. Most countries had indeed adopted regulations of this kind yet only a few contented themselves with them. The 1887 British law is representative of the "milder" type of intervention requiring only that margarine be properly labelled and that manufacturing establishments be registered and inspected. The Dutch legislation of 1908 was similar with the supplementary clause that margarine should not be manufactured, stored or sold in the same place as butter.³

Many countries have tried to prohibit the yellow coloring for margarine. Such a policy could not hide as easily its protectionist motive behind public health and consumer protection concerns. In the last two decades of the nineteenth century, Russia, France, Denmark, New Zealand and Australia passed legislation prohibiting the coloring of margarine. In the U.S., the story was complicated by the mixture of federal and state policies. State governments were the first to act and by 1886, no less than 22 states had some margarine legislation, 17 regulating labelling and packaging and seven - Maine, New

York, Pennsylvania, Michigan, Wisconsin, Minnesota, and Ohio - downright prohibiting the manufacture and sale of oleomargarine. Their prohibition bills often remained dead letter because there was no provision for enforcement and because the Supreme Court declared them unconstitutional in so far as they interfered with interstate commerce. This led the dairy lobby to turn to the federal government.

After "more thorough discussion and more reckless opposition than had been applied to any measure in Congress in the ten previous years,"⁴ the federal government passed in 1886 a legislation which introduced a manufacturing tax of 2 cents a pound and required quite expensive annual licenses for manufacturers, wholesalers and retailers of margarine (\$600, \$480, and \$48, respectively). There was much evasion of licenses particularly among retailers who continued to sell margarine for butter whenever they could. The dairy interests continued to lobby the government to bring margarine under control. State assemblies responded and by 1900, 32 states had passed legislation to prohibit the yellow coloring of margarine, the most bizarre certainly being Vermont in 1884 and West Virginia in 1891 requiring that it be colored pink. Finally, as in 1886, the 1902 federal amendment to the oleomargarine act was adopted only after a bitter fight in both houses of Congress. The manufacturing tax was raised to 10 cents a pound for colored margarine and lowered to 1/4 cent a pound for the uncolored product. The wholesalers and retailers' licenses were also reduced to \$200 and \$6 when they traded only in uncolored margarine. The American margarine policy is generally considered remarkably severe. But on a severity scale, Canada, and particularly the provinces of Quebec and Ontario [the two Canadian dairy regions], ranked, and still rank, higher.

A Century of Margarine Regulation in Quebec

From 1886 to 1949, the fabrication, sale and import of margarine was prohibited in Canada. The 1886 Prohibition Act did not cause as heated a debate as in the U.S. The House of Commons debated centered on the best way to protect the dairy producers and the consumers. In April, the MPs were talking about taxing and regulating margarine but in May, a bill prohibiting it was introduced, defended by the Ontario rural representatives⁵ and passed with very little opposition on June 1. There was in fact only one real opponent, Arthur Gillmor, MP for Charlotte, New Brunswick, who declared rather courageously that "If those who use oleomargarine were the electors of this country, you would not find so many to support this proposition." It was indeed true that the very poor, the most likely consumers of margarine at that time, could not vote since the franchise was reserved to property-owners.⁶

The story was quite different when margarine prohibition was reinstated in 1923, after being lifted as a war measure in 1917. By that time, as many opponents to the ban pointed out, the protectionist motive was the only one left, the dangerous nature of the product being impossible to defend in a world where Canada was alone to prohibit margarine. The anti-margarine were still, of course, the butter producers represented now by the National Dairying Association of Canada which comprised several provincial dairy associations. This time, Quebec was very present. A petition of 29,000 Quebec farmers and resolutions from 800 agricultural societies were sent to the Minister of Agriculture and most of the Quebec MPs argued and voted in favor of prohibition. Hence, one of the, Bouchard from Kamouraska, could say in 1923 that he would feel more at ease in his mother tongue but that he believes there are no more converts to be made among the French-Canadian members of the House.⁷

Women who just obtained the franchise in 1920 were at the foreground of the pro-margarine forces, either as consumers responsible for the household budget or as heads of women's organizations assuming their charitable duties towards the poor's children. The

groups who lobbied against the return of margarine prohibition were thus women's groups like the National Council of Women, the Montreal Women's Club, the Housewives' League, the Catholic Women's League, the Jewish

Council of Women, joined by the Great War Veteran's Association, the Retail Merchants Association of Canada, the Boards of Trade of many cities and the Salvation Army.

Faced with this sharp division in the public opinion and in the House of Commons, the Government played for time by permitting the manufacture and sale of margarine for one year at a time from 1919 to 1923. This made the political struggle to last six years and to fill hundreds of pages of Hansard with incredibly tedious and repetitive speeches about the importance of the dairy industry for Canada and about protectionism. The Mackenzie King government finally reinstated margarine prohibition in June 1923 arguing that the government had to respect the pledge made to the farmers during the war that the lift of the ban would last only as long as "abnormal conditions" prevailed. The vote crossed party lines and was 125 to 54.

In 1949, the 1886 federal margarine law was contested in the Supreme Court from three different sides. The government of Quebec argued that Trade regulation was a provincial right; a senator (Hayden) claimed that margarine prohibition favored one class at the expense of all others; and a consumers' representative, Margaret Hyndman, that it infringed on the civil rights of thousands of housemakers. The Supreme Court declared it unconstitutional on the grounds that margarine was no longer an injurious product against which the population had to be protected [a federal jurisdiction] and that it concerned not only interprovincial or international but also intraprovincial trade [a provincial jurisdiction].

From then on, margarine regulation became a provincial responsibility. Only three provinces took over: Ontario forbade (and still does) "butter yellow" margarine, while Quebec and Prince Edward Island prohibited its manufacturing and sale. As soon as the Supreme Court decision was known in December 1948, the Premier Ministre of Quebec, Maurice Duplessis, reassured the Quebec dairy farmers that his government would not let them down and warned margarine manufacturers not to make plans for Quebec too rapidly. Less than three months later, a bill with the revealing title "An Act to protect the dairy industry in the province" was adopted on March 10, 1949, almost unanimously (71 to 8). It authorized the government to adopt, upon the recommendation of the Minister of Agriculture, the regulations which he may deem expedient and just to prohibit or regulate the manufacture, sale and possession of margarine.

A week later, on March 17, an order in council prohibited the manufacture and sale of margarine in Quebec. The preamble is extremely instructive: there are nine "whereas", five of them concerned the dairy industry, two of them protection against foreign products, and two stated that the ban was also in the interest of the working class, the group most clearly disadvantaged by the measure. They may be worth quoting to illustrate the great amount of imagination legislators can show:

"Whereas it is necessary, in the interest of all classes of society, particularly in the interest of the working class, to maintain a suitable equilibrium between the rural population and the urban population;...

Whereas the weakening or disappearance of the dairy industry in the Province of Quebec would deprive the workmen and their children of a healthy and indispensable nourishment, one that is produced entirely in the Province of Quebec."

In 1961 margarine was finally legalized. At first sight, this legalization seems characteristic of the new spirit of the "Quiet Revolution" decade when Quebec society is said to have, belatedly but suddenly, stepped into the modern times. But in fact, it was the farmers' union (l'Union des cultivateurs catholiques) who asked the Quebec government to replace the ban by regulation. They felt that the 1949 ban had not been enforced and that

butter substitutes were sold as "spreads" in Quebec as freely as in the other provinces, even more freely than in Ontario, since these "spreads" were colored yellow.⁸ The new Liberal government complied with the 1961 Dairy Substitutes Act which regulated licensing, labelling, coloring, and advertising. As in Ontario, margarine could not be sold yellow with the result that manufacturers supplied a coloring capsule to the consumers to color it themselves, something housemakers did not seem to have appreciated much.

In 1969, the Dairy Products Act and the Dairy Product Substitutes Act were merged and the clause on coloring was relegated to an order-in-council in the Gazette officielle de Quebec, reducing the visibility of the anti-color regulation. It remained there until December when it quietly disappeared unnoticed. The 1987 new ban on yellow margarine ended the brief 15-year interval of total freedom since 1886. Once again, there was a violent debate between the farmers' organization (the Union des producteurs agricoles (ex-UCC) and the Federation des producteurs laitiers du Quebec) and the consumers' association and the margarine producers. Once again, the dairy interests won it. Quebec and Ontario might now be, so far as we know, the only places left in the world still regulating the color of margarine.

A Public Choice Explanation

Why was Quebec regulation so restrictive? Some may be tempted spontaneously by the cultural explanation, for instance the typical "protective impulse" of the Canadians or the "deeper-rural roots" of the Quebecers. The economic theory of regulation, however, suggests an interesting alternative. When they feel margarine is a serious rival, all dairy producers, in Wisconsin, New Zealand or Quebec, want governments to protect them. Whether they are successful depends on the opponents they are facing. If only consumers and margarine producers oppose margarine regulation, as in Quebec, the dairy lobby prospects are very good. In the public choice literature, this is a straightforward case of a policy with concentrated benefits and diffuse costs.

When other domestic agricultural interests are involved, the outcome may be quite different. The U.S. case is the best example. Until the First World War, most margarine was produced from beef fat (oleo) and some of it from cottonseed oil. This explains why the Western livestock interests and the Southern cotton growers were against the "federalization" of the prohibition already in force in the seven northern dairy states. They openly fought the 1886 federal margarine legislation and its 1902 amendment. During the war, the scarcity of animal raw materials and the new oil hardening process introduced from Europe in 1915 led to the use of coconut oil which accounted for 75% of the oil used in the fabrication of U.S. margarine in 1932. As a result, livestock farmers and cotton growers lost interest in the margarine industry which had become "un-American."⁹ Federal and state legislation against margarine grew in number and severity during this interwar period. Margarine regained political support from the Midwest farmers with the use of soya bean which came to represent by the 1960's close to 75% of the total raw material used in margarine. In 1949, the House of Representatives repealed the discriminatory tax on yellow margarine, although only by a slim majority of 152 to 140 votes. The state ban on coloring began to disappear in the 1940s, the last one being Wisconsin in 1967.

But cannot we say this about Canada as well? What about the cattle producers of the Western provinces and the Ontario soya bean producers? In the 1886 debate, the Western interests could not be present for the simple reason that the Prairies did not "exist" as a political entity. They were just beginning to develop and Manitoba was the only Prairie province with a mere 3% of the Canadian population at the 1891 Census. The Western MPs participated in the 1920s debate over the reestablishment of the ban and they were generally not anti-margarine. Heick (1991:32) illustrates well this point with his story of the fate of

the anti-margarine resolutions submitted to the annual meeting of the United Farmers of Alberta and of the Saskatchewan Grain Growers' Association: no one appeared in its support and it was "not dealt with."

The Ontario case is more complex. There were, and still are, both dairy producers and soya bean growers and the latter sell oil to the margarine industry and oilcakes to feed the cows. This perhaps explains why Ontario in 1949 adopted a "middle of the road" legislation instead of replacing the federal prohibition by one of its own. Newspapers reported in January 1949 that the 400 delegates to the meeting of the Ontario Federation of Agriculture disapproved the coming of margarine but felt it was inevitable. They were not asking that the Ontario government prohibit, but only that it be regulated.

In conclusion, the differences in the political landscape of Canada and U.S. have to be taken into account. In Canada, it just happened that the two dairy regions, Quebec and Ontario, were the two provinces with the most political clout.

¹House of Commons, Debates, 1918, p. 307.

²This account of the dietary needs which led to the invention of margarine is taken from Hoffmann (1969:11-13). To illustrate this point, Hoffmann gives some figures for Hamburg where 1 kg of butter cost 1.32 marks in 1850 and 2.44 marks in 1869 while a Prussian miner received in 1870 a wage of 2.59 marks for one shift.

³The information on margarine legislation outside North America is taken from Van Stuyvenberg (1969:290-318) and in the U.S. from Snodgrass (1930:28-116).

⁴As stated by the chairman of the House Committee on Agriculture. See Snodgrass (1930:33.)

⁵Hence, Heick's judicious choice of title "What Ontario wants, Canada gets..." for his 1988 article on the 1886 margarine debate.

⁶Gillmor in the House of Commons, Debates 1886, p. 552. The minimal value of property was about \$400 at the time of Confederation and was progressively reduced to allow retired people and well-off tenants to vote. As women did not have the franchise before 1920, voters represented only 15 to 25 percent of the population until the First World War. Bernard (1982:168).

⁷House of Commons, Debates 1923, p. 3562.

⁸The resolution adopted at the UCC's annual meeting in the fall of 1960 was reproduced next to the text of the Dairy Substitutes Act in the farmer union's weekly, *La terre de chez nous* of May 31, 1961, p. 6.

⁹This information comes from Van Stuyvenberg (1969:288). He argues without substantiating it that they went even further and formed a united agricultural front with the dairy farmers against the margarine industry.

SESSION FOUR - HISTORICAL ISSUES IN LABOR ECONOMICS
SUNDAY, 10:15 A.M.

**A LONGITUDINAL ANALYSIS OF THE SETTLEMENT PATTERNS,
OCCUPATIONAL MOBILITY AND WEALTH ACCUMULATION OF
EUROPEAN IMMIGRANTS TO THE U.S., 1840-1860**

Joseph P. Ferrie
Northwestern University

I. Introduction

The period of adjustment undergone by immigrants is a crucial part of the process of immigration. Though this process is interesting in its own right, as an indication of an economy's ability and willingness to absorb and embrace new members from vastly different backgrounds, it is important as well as a determinant of subsequent immigration: migrants whose expectations have not been met may discourage others from migrating. At the same time, the adjustment of immigrants to a new economy may postpone the benefit to the economy of the receipt of millions of working-age, often highly-skilled individuals.

For the U.S., the question of the extent of adjustment experienced by immigrants has received a great deal of attention. But few answers have been available for the early nineteenth century when the first great waves of immigration occurred. Though millions of European immigrants arrived in the years after the Napoleonic Wars and before the American Civil War, no data suitable for a large-scale analysis of the process of immigrant adjustment have been available for this period.

The decennial federal censuses of population -- the best source of economic and demographic information on America's population -- suffer from three shortcomings in addressing this issue: 1) the census did not ask immigrant respondents to state their year of arrival until 1890; 2) the census provides no sense of immigrants' circumstances in Europe or at arrival (e.g. their occupation); and 3) each census is a snap-shot of the population at a single point in time, while many questions of adjustment can best be addressed with data which are longitudinal.

To overcome these difficulties with the census and yet still have a body of data broad enough to allow generalizations to be made about the experiences of a large number of immigrants, I have produced a sample of more than two thousand male immigrants who were followed from their arrival in the U.S. (as documented in passenger ship arrival records) until they turned up in the manuscript schedules of the 1850 or 1860 federal census of population; more than five hundred immigrants in the sample appeared in all three sources, providing a record of their experiences which extends from their arrival for up to twenty years.

In this essay, I will use these data to explore three aspects of the process of immigrant adjustment: patterns of settlement, occupational mobility, and wealth accumulation. The questions I will address are: 1) how extensive were the changes in location and occupation made by immigrants, and how likely were immigrants to make such changes as their time in the U.S. increased?; and 2) what was the impact of those changes, as measured by their effect on wealth accumulation?

Few of the migrants studied here simply "uprooted" in the sense of the trauma, confusion, and paralysis described by Handlin -- a clear process of adjustment is evident beginning from the time of arrival. Many appear to have been "transplanted" instead -- arriving with numerous handicaps (a lack of skills, capital, or knowledge of English), but acting after their arrival in ways consistent with attempting to reduce the impact of those

deficiencies. The data suggest that immigrants continued to search for areas in which their greatest promise lay from arrival right through at least the end of the twenty years covered by the sample.

II. Patterns of Settlement

Where immigrants were located by the time of the 1850 and 1860 censuses is shown in Table 1. [Notes, references, tables, and figures will be provided at the session] Seventy-two percent of the immigrants in the sample in 1850 were located in the Northeast, and more than half were found in just three states -- New York, Pennsylvania, and Ohio. Smaller numbers were found in New England and the states closest to the frontier, while very few were found in the South.

Immigrants were clearly over-represented in cities: though more than 87% of the total U.S. population in 1850 lived in places of fewer than ten thousand people, only just over half of the immigrants in the sample were found in such places in 1850, and only 44% were found there in 1860. Part of these disparities might result from the fact that many cities were transportation hubs through which immigrants had to pass before they could reach the interior where most natives were located. Immigrants were indeed over-represented in New York City, Philadelphia, Cincinnati, and Buffalo, all of which served as gateways to the interior. But this does not account for the concentration of immigrants in smaller cities, places with fewer than 50,000 people.

Even the earliest arrivals in the sample were more concentrated in cities than were natives, which suggests that immigrants were attracted to urban centers for reasons other than the fact that they had to pass through them to get anywhere else. The attractions of cities included the high level of average wages, ready sources of employment for the unskilled and, in many cases, the presence of large numbers of one's countrymen -- an important source of both potential employers and potential customers.

Immigrants who arrived at the end of the 1840s were no less likely than those who arrived at the start of the 1840s to be found in 1850 in locations far removed from New York, their point of entry. At their arrival in New York, immigrants faced a variety of relatively inexpensive routes to the interior of the U.S., which many had arranged to employ even before their departure from Europe. The total fare from Europe to Buffalo, Cleveland, or Detroit was probably only one fourth greater than the \$35 fare to New York city alone, with the total fare to Chicago and interior points such as Cincinnati, Pittsburgh, and Louisville only a third more than the fare to New York alone.

Though early and recent arrivals were found in many of the same locations in 1850, it was the recent arrivals who were more likely to relocate between 1850 and 1860. Table 2 shows that thirty percent of all immigrants in the sample linked from 1850 to 1860 changed region between 1850 and 1860, with net movement being from east to west: 15% moved to more western locations, while 9% moved from western locations back into the Northeast. Even among the earliest arrivals, movement between 1850 and 1860 was not negligible. Of the one hundred immigrants in the sample who arrived in 1840 and 1841, nineteen changed regions between 1850 and 1860 -- the process of adjustment clearly continuing into the second decade of these immigrants' lives in the U.S. But the magnitude of this change is dwarfed by change among the most recent arrivals: of the 105 arrivals in 1849 and 1850, 50 had changed locations between 1850 and 1860.

It thus appears that immigrants moved quickly from their point of entry to their 1850 location, with the vast majority (seventy percent) remaining in the same region over the period 1850 to 1860. But nearly thirty percent seem also to have found an 1850 location rapidly, but only over the course of several years found a location in which they were likely to remain. Changes in location between 1850 and 1860 could have resulted from several circumstances: receipt of information regarding superior locations, spending time at one location to accumulate the human or financial capital which would permit a move to a better location, or simply a change in fortune which permits or requires such a change.

The other important change in location which occurred in the sample between 1850 and 1860 was net migration from rural to urban locations. Nearly twice as many migrants went from rural to urban places as from urban to rural places. Those most likely to make either sort of move were the most recent arrivals. But the average age of those going from rural to urban places was more than two years less than that of those who remained in rural places, while urban-to-rural movers were no different in age from those who remained in urban areas.

Since younger workers had probably made less of an investment in their occupation by 1850, while they had a longer time over which to reap the net benefits of changing occupations, it is possible that those rural-to-urban moves represented changes in occupation as well, while urban-to-rural moves were immigrants merely passing through on their way to a rural occupation in which they expected to remain. In fact, of the 63 immigrants who made rural-to-urban moves, 40 (64%) changed occupation as well as location, while only 40% of the rest of the entire sample changed occupation. Among those making urban-to-rural moves, the same percentage changed occupation (43%) as among the rest of the sample. A more detailed consideration of occupational mobility follows.

III. Occupational Mobility

The distribution of 1850 and 1860 occupations in the sample shown in Table 3 is more concentrated in white collar and skilled blue collar occupations than the total U.S. population: 44 percent of the sample was in these occupations by 1850, while only 30 percent of the total U.S. population was so employed. Immigrants were correspondingly under-represented as farmers and laborers (56 percent versus 63 percent for the total U.S.). This difference between the sample distribution and the U.S. distribution widened between 1850 and 1860.

Most immigrants changed occupation at least once in the years after their arrival. Table 4 reports these findings. More than half changed occupation between arrival and either 1850 or 1860, with just over 40% changing between 1850 and 1860. For example, 9% of those in the sample described themselves as farmers at arrival and as laborers in the 1850 census. Of those moving from farmer to laborer by 1850, 63% had moved back up out of the laboring class by 1860. Overall, 16% of the immigrants in the sample arrived as farmers but had changed occupation by 1850. Of these, only 9% had returned to farming by 1860, while 67% had entered white collar, craft, or mining occupations.

Although there appeared to be considerable downward occupational mobility among these farmers at arrival who were found in occupations associated with lower wealth by 1850 or 1860, at least some of it may have been the result of a planned change of occupation, rather than an inability to re-attain a previous occupation. Farmers who were temporarily unable to obtain their own farms to work would have been most likely to show up in the census as farm laborers, rather than undertaking the investment necessary to convert to a white collar or craft occupation. At the same time, those who left farming for another occupation, both between arrival and 1850 and between 1850 and 1860, were more than five years younger on average than those who remained in farming.

That the decision to leave farming was made by just those immigrants who would have been most likely to make a decision to do so permanently (younger farmers), while

those changes were not the sort we would expect to see made by those hoping to remain in farming suggests that leaving farming was for at least some the result of a decision to forsake permanently farming for another occupation rather than a temporary expedient before an eventual return to farming. The next section considers the impact of such changes on wealth.

IV. Wealth Accumulation

In his landmark study of 19th century wealth-holding patterns, Soltow (1975) found a concave relationship between age and wealth. A large literature in economics describes a variety of mechanisms which could account for such a relationship. But Soltow also discovered marked differences in wealth between the native and foreign born, as well as differences by country of origin among the foreign born. He suggested that these differences might have resulted from differences in immigrants' time of arrival, and in their ability in the years since arrival to overcome barriers posed by language, prejudice, and low levels of skill at arrival. But, lacking any information on immigrants at their arrival, he was unable to assess the importance of these factors.

Tables 5 and 6 show the results of regressions on the natural logarithm of real wealth in 1850 and 1860 for all immigrants located in either 1850 or 1860. The explanatory variables include age and its square, fixed factors such as nationality (over which the immigrant had no control), variable factors such as location and occupation (which immigrants could alter), combinations of fixed and variable factors such as the impact of changes from occupation at arrival (which the immigrant could not control in 1850 or 1860) to occupation in 1850 (which he could control), and the independent effect of time since arrival.

The results display the expected concave age-wealth relationship: wealth rises with age at a diminishing rate in both years. But the age-wealth profile is both steeper and later-peaking in 1860. At age 30, for example, wealth is growing at 9.3% in 1860 but at only 7.2% in 1850; the 1860 profile achieves a maximum at age 59.2, while the 1850 profile peaks at 50.9 years. The Irish were 30% less wealthy than the British and Germans in 1850, but this difference was halved and reduced to statistical insignificance by 1860.

The results show much higher wealth for those in the Northwest states in both years and in the North Central states in 1860, and for those outside urban places. Those located in the Northwest, for example, were nearly two and a half times wealthier in 1860 than identical individuals located in the Northeast. Such differences probably resulted from the superior rate of return on investment in land in the west, as well as from a return to the riskiness of locating there. Again, though, there are pronounced differences between 1850 and 1860. The advantages of those in the Northwest and North Central states, as well as the disadvantage of those in urban places, are all greater in 1860 than they had been in 1850.

The results also display a clear hierarchy of wealth by occupation, much like that observed in other studies of nineteenth century wealth. Farmers are clearly the wealthiest: they own on average 20 times the wealth of laborers in 1850. Those in white collar occupations are one and a third times wealthier and craftsmen and miners are 91% wealthier than laborers in 1850. But again, the advantages of particular occupations are considerably greater in 1860 than they had been in 1850. Farmers were nearly 60 times wealthier than laborers by 1860, while the gaps between white collar workers and laborers and between craftsmen and laborers widened correspondingly.

The effect of occupational change since arrival shows the role of previous occupation (and presumably previous ability to accumulate wealth) on the level of wealth in 1850 and 1860. Craftsmen and miners in 1850 who were described as farmers in the passenger ship lists were nearly 50% wealthier than craftsmen and miners who were craftsmen and miners in the ship lists. Since farmers possessed the greatest wealth (both real and total) of any occupation in the 1850 and 1860 census, this suggests that immigrants

who arrived as farmers may have had more wealth at arrival than those in other occupations, and that some of this advantage persisted even after a change in occupation. Laborers who had risen into the ranks of craftsmen and miners were 53% less wealthy than craftsmen and miners who had not changed occupation between the passenger ship lists and the 1850 census. This might be the result of the persistence -- even after occupational change -- of the disadvantage in wealth of those who arrived as laborers. By 1860, though, these effects were greatly reduced in statistical significance.

Finally, the return to duration after these other factors have been taken into account is both substantively and statistically significant in both years, though it smaller in 1860: wealth rose by 8.2% for every year since arrival in 1850, but by only 6.0% per year in 1860. It is possible that the return to duration in the U.S. is an illusion. If there was a secular decline over the 1840s in the wealth of all immigrants at their arrival -- a cohort effect -- then we might in considering a single cross-section of individuals mistakenly attribute the greater wealth of earlier arrivals to their greater time in the U.S., when in fact they were merely wealthier on average when they got here than later arrivals.

However, the fragmentary evidence which can be mustered on this point suggests that, if there was any cohort effect, it actually works to make the observed duration effect smaller than the true effect. The average wealth of German emigrants at their departure from Germany, as recorded by officials of their principalities, shows a slight rise over the 1840s, so the actual rate of wealth accumulation with duration in the U.S. for Germans at least is probably greater than the rate implied by Tables 5 and 6.

How can we account for the magnitude of this duration effect? One possibility which we can immediately discount is that the duration effect represents mainly the acquisition of language skills by non-English speaking immigrants. When regressions are estimated separately by country of origin, the Germans do indeed have a larger duration effect than the Irish, but the British have the greatest effect of all. The return to duration in the U.S. might also represent the acquisition of skills and knowledge other than language specific to the U.S. economy.

Another possibility is that the effect of duration in the U.S. is capturing general growth in the value of real estate or general economic growth over the 1840s and 1850s. Though there are no figures on the rate of growth of the value of real estate in the U.S. between 1840 and 1850, the corresponding figures for the next decade suggest that per capita real wealth grew at perhaps 5% per year between 1850 and 1860. Though the return to duration was higher than this average in some places -- as high as 8.3% per year between 1850 and 1860 in Chicago -- nearly two thirds of the immigrants in the sample were located in the Northeast where such high rates of return to duration in specific local economies were probably rare.

Much of the return to duration may simply represent the higher average level of per capita income in the U.S. relative to Europe. The fact that an immigrant who has been in the U.S. for only a year is 8% wealthier than an otherwise identical immigrant who has just arrived might result from the fact that the immigrant in the U.S. for a year has had the opportunity to earn for that year an income well above that earned in Europe by the more recent arrival in the same year.

Another result of duration in the U.S. is more subtle, though of perhaps greater magnitude: the increase in the returns to the locations and occupations into which immigrants moved after their arrival and the increase in the age at which wealth peaked. The increased advantage to those in particular locations and occupations is consistent with immigrants better matching themselves to places and jobs in the years after arrival. By 1860, a larger number of successful matches had been made than by 1850. The reduction in the wealth disadvantage of the Irish between 1850 and 1860, after accounting for all other observed characteristics, is also consistent with more successful matching: having come from Ireland with few skills and little capital may have been less of a handicap after a decade

in which one could acquire additional skills or learn where best to put one's existing skills to work.

The later age peak in wealth in 1860 is consistent with immigrants over the course of their time in the U.S. economy being able to take better advantage of their comparative advantage in particular pursuits. For example, laborers relied more on simple physical strength than those with greater skill (craftsmen and those in white collar occupations) or those whose income was earned with inputs complementary to physical strength (farmers who used land, animals, and farm machinery). The peak of the age-earnings profile for laborers would have occurred earlier than for workers with greater skill or other inputs to employ, and the age-wealth profile of skilled workers or farmers would have occurred later as a result. More immigrants had been able to move into occupations which made use of more than physical strength by 1860, allowing them to postpone the age at which earnings and wealth peaked.

To see whether immigrants making changes in location or occupation (two of the most important choice variables in the regressions) are indeed responsible for shifting out the peak of the age-wealth profile, the regressions in Tables 5 and 6 were re-estimated, on the sample of immigrants located in both 1850 and 1860. The estimated age-wealth profile for the entire linked sample is shown in Figure 1. The peak rises from 48 years in 1850 to 62 years in 1860.

Figure 2 shows the result when immigrants who left the Northeast (most of whom relocated to the North Central and Northwest states) between 1850 and 1860 are excluded from the sample. The result is a much smaller shift in the peak of the age-wealth profile, moving from 47 years to 59 years, a shift of only 12 years rather than the 14 year shift when those who left the Northeast were included. Figure 3 does the same, but excludes those who entered farming between 1850 and 1860; the result is similar -- the peak shifts only 11 years. Finally, Figure 4 shows the result when those relocating from urban to rural places between 1850 and 1860 are excluded: the peak of the profile shifts only nine years.

The extent of adjustment resulting from such changes is probably a lower bound on the amount of adjustment actually experienced. The results in Figures 1 through 4 include the impact on wealth of changes in location or occupation, but exclude the impact of changes within locations or occupations -- finding a more favorable location within the Northeast or a better way to utilize one's existing skills, for example. The effect of such hidden adjustments is another component of the 6 to 8% return to duration.

V. Conclusions

Several of the same trends in the relationship between the age-wealth profile and duration in the economy uncovered here have been observed in the development of the Utah economy and the economies of several western states and counties. (Pope 1989 and Soltow 1975) In those cases where age-wealth profiles can be traced over a period of decades for an economy, an increasingly mature economy consistently displays a higher and later-peaking profile than had the same economy earlier in its development.

The explanation offered for those changes -- that in an increasingly mature economy, individuals may have had more opportunities to take advantage of their skills and thereby change the shape and peak of the age-earnings and age-wealth profiles -- might account for much of the difference between the profiles in 1850 and 1860. Immigrants who had spent a longer time in the U.S. (the 1860 sample) might be like individuals in an economy at a more mature stage of development, while recent arrivals (the 1850 sample), like individuals in a newly-formed frontier economy, had yet to sort themselves into the occupations and locations in which their skills could earn the greatest return.

THE ECONOMIC RATIONALE OF APPRENTICESHIP

Bernard Elbaum

Nirvikar Singh

University of California, Santa Cruz

Apprenticeship, a pervasive institution within medieval Europe, persists today in various modified forms as an important element in different national systems of education and training. The persistence of apprenticeship has generated a variety of theories regarding its underlying economic rationale. According to one school of thought, apprenticeship results from monopolistic restrictions imposed by trade unions, employers' associations, and/or the government. From another perspective, apprenticeship offers efficiency benefits that either explain why unions, employers, and the government could find it in their joint interest to maintain apprenticeship regulations, or more directly, why apprenticeship managed to persist of its own momentum.

Although the workings of apprenticeship, both in the present and historically, have been much discussed, they have never been the subject of formal economic analysis. The present paper offers a first attempt at a formal economic model of apprenticeship which seeks to highlight the efficiency benefits that may provide at least part of its underlying rationale.

Our model is based on British apprenticeship practices of the late nineteenth and early twentieth century, and draws upon related historical research. The British case is of particular interest because its apprenticeship practices, medieval in origin, persisted into modern times amidst broadly competitive labor markets and laissez faire government. British laws requiring apprenticeship to practice a skilled trade and enjoy the rights of freemen were abolished respectively by repeal of the Statute of Artificers in 1814 and passage of the Municipal Corporation Act of 1835. Subsequently, apprenticeship flourished within Britain into the post-World War II era without benefit of active public promotion.

Because apprenticeship practices in different times and places tend to have at least some features in common, our model may also be of wider applicability. We will, in particular, try to apply it in order to understand why British apprenticeship practices, though transplanted into colonial America, largely failed to persist within the United States.

Apprenticeship, we suggest, was a device which allowed financially constrained youths to exchange indentured labor services for employer financing of training investments. To the extent that it permitted greater investment in training, apprenticeship afforded an efficiency improvement over the alternative of an unstructured labor market. But apprenticeship was only a "second best" arrangement.

The firm that employed apprentices was constrained to pay an initial wage high enough to meet the subsistence needs of the apprentice's family. After investing in training, since the craft skills involved were potentially transferable to other employers, the firm was further constrained to pay a wage high enough to deter quits that would entail the loss of its investment. As a result, apprentices tended to earn quasi-rents. Because enterprise investment responds in general only to the returns that the firm itself can capture, the economic logic of apprenticeship tended to lead, from a societal standpoint, to underinvestment in training.

Central to the logic of apprenticeship, according to our model, is the indenture agreement, through which the worker and employer jointly commit to a fixed term of employment as well as to the provision of training. In Britain a variety of mechanisms helped ensure that indenture agreements would be honored, and that apprentices, in particular, would complete their stipulated term of indenture. Indenture agreements, when written, were legally binding, and even when unwritten, as they usually were, possessed the quasi-legal legitimacy of long-standing custom. In addition, employers' associations as well as trade unions pressed firms to require certification of indenture completion for entry into skilled

jobs. Finally, and we suggest, most importantly, firms regarded indenture completion as verifying the worthiness of youths for skilled employment--behavior which we describe formally through a model of skill certification.

Our model also suggests that certification requirements were essential in order to forestall an apprentice runaway problem that could lead to the collapse of the apprenticeship system. The decline of apprenticeship in the United States appears to be a case in point. In this rapidly expanding settler society, no certification was needed to obtain well paid work in the skilled trades. Chronic scarcity of skilled labor led employers to the common practice of filling skilled jobs by up-grading less skilled workers. Mobility--both geographic and occupational--was high, especially among youths, and tended to undermine the enforcement of apprenticeship indenture. Although the powers of the law, trade unions, and employers associations were invoked to enforce apprenticeship regulation, they proved unable to stem the tide that erased all but a small trace of British apprenticeship traditions from the American continent.

The paper proceeds by describing British apprenticeship practices in outline, emphasizing the evidence that supports the two key assumptions of our model: (1.) apprenticed workers were subject to a subsistence constraint; and (2.) the possibility of apprentice quits constrained the pay that firms offered in the latter stage of indenture.

We next present our formal model. For simplicity, we assume that an individual's work life consists of n periods, of which the first two may be occupied by apprenticeship. Firms may invest in training only during the initial period of apprentice employment. For apprenticeship to be profitable the apprentice quit rate must be low enough in the second period for firms to recoup their initial training investment.

We consider three models that are ordered in terms of the scope of the factors they seek to explain. We begin by assuming a fixed quit rate that is given exogenously. The propensity to quit may in this context be viewed as being equal to the proportion of employers willing to hire runaway apprentices at their marginal product--a proportion determined by some form of collective regulation on the part of unions, employers' associations, or the government.

We then analyze an alternative model, in which an apprentice's decision to quit or complete his indenture depends on the wage policy of his employers, and in particular, on how the second period wage compares with an exogenous distribution of external wage offers. In this model, the firm optimizes with respect to its choice of a second period wage.

Finally, we present a certification model in which the distribution of external wage offers, and hence of apprentice quits, is determined endogenously. In this model apprenticeship completion may be viewed as playing much the same kind of a role as a school degree today. By assumption, the firm can only discern the productivity of workers who lack skill certification through a costly and time-consuming process of work supervision. As a result, the firm is only willing to hire such workers at a pay rate which corresponds to its initial probabilistic estimate of their productivity. Apprentices are in turn deterred from running away since if they do they cannot easily be distinguished from workers who lack systematic training, or from individuals who dropped out of apprenticeship because they were unsuited to the work.

After deriving our formal results, we review evidence on British apprenticeship which suggests that collective regulation of indenture agreements was relatively ineffective in engineering and building--by far the largest sectors of apprentice employment--and draw the implication that enforcement of indenture obligations here depended on the preference by individual employers to hire craftsmen of certified skill attainment. We then proceed to examine the decline of apprenticeship in the United States.

THE DECLINE IN BLACK TEENAGE LABOR FORCE PARTICIPATION IN THE SOUTH, 1900-1970: THE ROLE OF SCHOOLING

Robert A. Margo T. Aldrich Finegan
Vanderbilt University

In 1950, approximately 61 percent of southern black teenagers, ages 16 to 19, were participating in the labor force. By 1970 their participation rate had declined to 34 percent, or by 27 percentage points in two decades. Perhaps the most widely accepted explanation of the post-1950 decline in black teen participation emphasizes decade-specific shifts in labor demand (John Cogan; Gavin Wright). Between 1950 and 1960 mechanization of agriculture in the South accelerated significantly. In particular, the introduction of the mechanical cotton picker disproportionately reduced the demand for the labor of black male teenagers, who were concentrated in cotton production. Black teens displaced from southern cotton farms did not find ready work in the South's non-farm economy. Their failure to find non-farm jobs, according to Cogan, was caused by the federal minimum wage, primarily through increases in its coverage in the 1960s. The combination of mechanization and the minimum wage drastically reduced job opportunities for black male teens in the South, resulting in the sharp decline in labor force participation between 1950 and 1970.

We argue that the quantitative importance of demand-side shocks in accounting for the post-1950 decline in black teen participation has been overstated. A significant portion of the post-1950 decline appears to have been a continuation of a downward trend whose origins can be dated to the early 1900s. In turn, the long-term downward trend in participation was the mirror image of a long-term upward trend in black school enrollment. Each successive generation of southern black males was more likely to be in school between the ages of 16 and 19 than out of school and in the labor force. Further, the labor force participation rate of black male teens simultaneously enrolled in school declined over time, especially before 1950. Because the factors that were responsible for the long-run increases in black schooling had not yet run their course by mid-century, a post-1950s decline in labor force participation of black male teens in the South would have happened anyway, even if mechanization in cotton agriculture had not occurred in the 1950s and had the level and coverage of the minimum wage not been increased in the 1960s.

Census data provide the basis for constructing estimates of pre-1950 labor force participation rates of black male teens in the South. At the turn of the century fully 86 percent of southern black male teens were participating in the labor force. Their participation rate fell by 10 percentage points between 1900 and 1920, and by an additional 15 percentage points between 1920 and 1950 -- or an average rate of decline of almost 5 points per decade between 1900 and 1950. Fully 48 percent of the decline in participation between 1900 and 1970 took place before 1950.

Cogan attributed the post-1950 decline in black teen participation to the negative effects of mechanization on labor demand in cotton agriculture and of the minimum wage on labor demand off the farm. Between 1950 and 1970 the number of man-hours required to produce a bale of cotton declined by 84 hours, principally due to the widespread adoption of the mechanical cotton picker. Because cotton production remained constant over the period, the decline in labor requirements translated into a one-for-one reduction in labor demand. Cogan claimed that virtually the entire decline in employment of black teens in the South between 1950 and 1970 could be explained by the decline in their employment in agriculture. Although some black teens found jobs in manufacturing, their employment growth in retail trade slowed considerably in the 1960s, precisely when coverage of the minimum wage was extended over this industry. Based on state-level regressions, Cogan

concluded that technical change and the minimum wage were equally important in explaining the decline in participation between 1950 and 1970.

Whatever the merit of Cogan's analysis of post-1950 trends, census data demonstrate that participation rates of southern black male teens were already falling between 1900 and 1950. Indeed, had post-1950s participation rates continued to decrease at the average decadal rate experienced between 1900 and 1950 (4.96 percentage points), the predicted participation rate in 1970 would have been 51 percent. The difference between this predicted 1970 rate and the actual 1950 rate accounts for 37 percent ($= 9.9/27$) of the decline in participation between 1950 and 1970. If the predicted 1970 rate is calculated on the basis of the average decadal rate of decline experienced between 1930 and 1950, the percent explained increases to 49 percent. Clearly, the rate of decline in black teen participation accelerated after 1950, but a reduction in participation could have been predicted (and was, by John Durand) on the basis of pre-1950 trends.

The diffusion of the mechanical cotton picker cannot explain the decline in black teen participation prior to 1950, because less than 1 percent of southern cotton acreage was mechanically harvested in 1950. Might agricultural mechanization or the minimum wage account for any of the pre-1950 decline in participation rates? Consider first the decline in participation before 1930. The minimum wage cannot possibly explain the pre-1930 decline in participation because minimum wage legislation was enacted in 1938. The principal type of mechanization occurring in American agriculture prior to World War Two was the diffusion of tractors. The diffusion of tractors was far slower in the South than elsewhere in the country. In 1930, only 4 percent of southern farms used tractors. Yet the participation rate of black teens fell by 12 percentage points between 1900 and 1920, before virtually any tractors had been introduced into southern agriculture.

In the case of cotton production before the diffusion of mechanical cotton pickers, tractors did not reduce labor requirements in farm tasks that were intensive in the use of labor. Further, the diffusion of tractors in cotton production was comparatively slow, yet it was this sector of southern agriculture that was relatively intensive in the use of black teen labor. Nevertheless, it is possible that tractorization or the minimum wage might account for some of the decline in participation rates between 1930 and 1950. The proportion of southern farms using tractors rose to 26 percent by mid-century. Recent research demonstrates that tractorization was associated with a decrease in sharecropping and other forms of tenancy and with a redistribution of cotton production to the Southwest and California. Because tenant labor was frequently supplied in family groups, the decline in tenancy may have reduced, *pari passu*, the demand for the labor of black male teens on the farm, as might have the spatial reallocation of cotton production. Estimates of coverage rates of the minimum wage indicate that about 18 percent of employed southern black teens in 1950 were in covered employment. The establishment of the federal minimum wage in 1938 may have reduced job opportunities for black teens even before 1950, and hence made labor force participation less likely.

To evaluate the pre-1950 impact of tractorization and the minimum wage, we estimated a pooled time-series cross-section fixed-effects regression using state-level data for 1930 and 1950. The dependent variable is the participation rate of black male teens; the independent variables are the proportion of farms using tractors, the coverage rate of the minimum wage, and dummy variables for year (1950) and for states (the state dummies are the fixed effects). The effects of tractorization and the minimum wage are negative as hypothesized, but are economically and statistically insignificant: together, tractorization and the minimum wage can explain about 10 percent of the 1930-50 decline in participation. Put another way, the likelihood a black male teen would participate in the labor force fell between 1930 and 1950 throughout the South. Variations across states in the rate of decline were essentially unrelated to the diffusion of tractors or the extent of coverage of the minimum wage.

Our analysis might be criticized, however, because of alterations in labor force statistics over the period. In 1940 the definition of the labor force changed from the "gainful worker" concept to the "labor force week" concept. As Durand and others have pointed out, the pre-1940 participation rates are not, strictly speaking, comparable with the post-1940 estimates. Two responses to this criticism can be made. First, changes in the definition of the labor force have little bearing on the finding of a pre-1940 decline in participation, because the gainful worker definition was used in 1900 and 1930. Second, Durand calculated adjustment ratios in an attempt to make the pre- and post-1940 participation rates comparable. The force of Durand's adjustments was to reduce the size of the labor force prior to 1940. Had the 1940 census question been asked in 1900 and 1930, participation rates in both years would have been lower.

Unfortunately, data limitations prevented Durand (and us) from estimating decade-specific adjustment ratios for southern black male teens, but he did estimate an adjustment ratio for all males ages 14 to 19 in 1930: 0.9756. Multiplying our estimate of the 1930 participation rate for southern black male teens by Durand's adjustment ratio produces an adjusted participation rate of 72.1 percent, compared with the measured participation rate of 73.9 percent. Although there is reason to believe that the true adjustment ratio for southern black male teens was higher than 0.9756, even if we assume an adjustment ratio equal to the lowest estimated by Durand for any group of males (0.8953, ages 75 and over) there still was a decline in black teen participation between 1930 and 1950. In sum, it is doubtful that adjusting for definitional changes would reverse the decline in labor force participation by southern black male teens before 1950.

If the minimum wage, agricultural mechanization, or biases in census data cannot explain the majority of the pre-1950 decline in black teen participation, which factors can? Some clues are by examining participation rates by farm-nonfarm and school enrollment status, along with the percent on farms and percent enrolled in school. At the turn of the century over 90 percent of southern black male teens not enrolled in school were participating in the labor force. Participation rates of teens enrolled in school were lower than participation rates of those not in school, but were still relatively high by comparison with the rates experienced after World War Two. Despite a lower participation rate among in-school non-farm teens, the overall participation rate did not differ between the farm and non-farm population because school enrollment was more common among farm teens.

In 1950, the participation rate of farm teens not enrolled in school was essentially the same as in 1900. The participation rate of farm teens enrolled in school, however, declined by 20 percentage points and the proportion enrolled increased by 21 percentage points, between 1900 and 1950. Among non-farm teens, the percent enrolled in school increased by 34 percentage points and the participation rate of in-school black teens declined by 22 percentage points between 1900 and 1950 but, unlike the farm population, the participation rate of nonfarm teens not enrolled in school decreased by 12 percentage points. As a consequence of these changes, the overall participation rate of nonfarm teens fell 33 percentage points between 1900 and 1950, almost three times the decline experienced by farm teens. In short, the overall participation rate for nonfarm teens fell more than the rate for farm teens because of a larger rise in nonfarm school enrollment and a larger fall in participation among nonfarm teens not in school.

School enrollment rates in both sectors continued to climb between 1950 and 1970, at a quicker pace than before 1950. Among farm teens, participation rates of those enrolled in school fell by 34 percentage points, but the participation rate of non-farm teens enrolled in school remained stable at slightly over 20 percent. In contrast to the pre-1950 period, participation rates of farm teens not in school declined sharply, by 23 percentage points between 1950 and 1970. Participation rates of nonfarm teens not in school continued to fall as they had before 1950, with the majority of decline concentrated in the 1950s. Thus,

reversing the patterns for the pre-1950 era, from 1950 to 1970 the overall participation rate for farm teens fell more than twice as much as the rate for nonfarm teens.

This analysis suggests that increasing school enrollments and reductions in labor force activity among those enrolled in school played a key role in generating the pre-1950 decline in black teen participation. The hypothesis is confirmed by counterfactual participation rates for 1900 and 1950 under various assumptions about school enrollment rates and participation conditional on enrollment status. Had enrollment rates (farm and nonfarm) in 1900 been equal to their 1950 values, the overall participation rate would have been 77 percent. Had enrollment rates and participation rates of in-school black teens in 1900 equalled their 1950 values, the overall participation rate would have been 68.2 percent. The difference between the actual 1900 participation rate and the latter counterfactual rate accounts for nearly three-quarters of the decline in participation between 1900 and 1950. By contrast changes in participation among those not enrolled were less important, explaining about 18 percent of the decrease in the overall participation rate before 1950.

Increasing school enrollments and decreasing participation rates of those enrolled in school continued to be important after 1950, but the relative significance of changes in participation rates of those not in school was greater after 1950 than before. Because nonfarm participation rates were lower than farm participation rates, the shift of population out of agriculture between 1950 and 1970 (35 percentage points) also helped to lower the overall participation rate.

The sharp decline after 1950 in participation rates of farm teens not in school is prima facie evidence that mechanization reduced the agricultural demand for teen labor. The shift of black labor out of agriculture, the increase in school enrollments, and the decrease in participation among those in school might also have been a response to technical change, except that downward trends in these variables began before the mechanization of cotton agriculture. What needs to be explained are the relatively low rate of school enrollment and high participation rate of black teens in school prior to 1950.

We hypothesize that changes before 1950 in school enrollment rates and in participation rates of in-school black teens were driven by increases in educational attainment of successive generations of southern black children. Historically black parents in the South were poor and heavily concentrated in rural areas. Because black parents were poor, because the returns to schooling were perceived to be low in agriculture, and because the labor of teenagers was valued on the family farm or in the nonfarm labor market, black children were not expected to remain in school at older ages. In 1910, for example, the peak age of black school enrollment in the South was age 11 (71 percent) and enrollment rates fell off sharply at age 14. The upshot was that southern blacks educated in the late nineteenth and early twentieth centuries typically completed very few years of schooling before reaching adulthood -- an average of about 5 years for the birth cohorts 1886 to 1905.

That completed years of schooling were so low can account for the low enrollment rate of black teens ca. 1900 but it cannot explain why the labor force participation of in-school black teens was so high in the early 20th century. A child whose education stopped at the sixth grade could, in theory, finish his schooling by age 12 or 13, and then enter the labor force. In fact, however, a significant fraction of southern black children suffered from high rates of age-in-grade retardation, and simply did not complete a given number of years of schooling in the customary amount of time. According to the 1940 census, for example, 47 percent of twelve year-old southern black males enrolled in school were attending the fourth grade or lower. Grade retardation in the pre-teen years had many causes, including adult poverty, and employment of young children in agriculture. But there is little doubt that a major factor retarding black educational achievement was the poor quality of segregated public schools attended by southern black children, especially the shorter school terms characteristic of the black schools.

As long as the demand for completed years of schooling was low, age-in-grade retardation at an early age was not a binding constraint. Completion of a given number of grades could be spread over several years and, if necessary, combined with labor force participation at later ages. Once the desired years of schooling reached modern norms, however, enrollment rates at later ages increased and participation of in-school black teens fell. A black teen struggling to finish his schooling with the sixth grade could work and still expect to achieve his goal before leaving home. But a black teen whose parents wished him to complete high school would, under the best of circumstances, normally graduate at age 18. He could ill afford to fall several grades behind, at an early or a later age, and his parents could best accommodate by keeping him out of the labor force during his teenage years. Over time, more black parents sought to have their children complete high school, and improvements in the quality of black schools lessened the extent of early age-in-grade retardation.

To determine the explanatory power of this hypothesis, we estimate the following two regressions:

$$(1) \text{ LFPR}_{ij} = a_{0i} + a_{1j} + a_2 \text{ELEM}_{ij} + e_{1ij}$$

$$(2) \text{ ENR}_{ij} = b_{0i} + b_{1j} + b_2 \text{ELEM}_{ij} + e_{2ij}$$

where:

LFPR: labor force participation rate of in-school teens
 ENR: percent enrolled
 ELEM: percent of enrolled black teens, ages 16-19, attending the elementary grades
 e: random error term

The subscript *i* refers to the urban, rural nonfarm, and rural farm sectors, and the subscript *j* refers to states. The dummy variables for sectors and for states control for a variety of factors that may have influenced participation and school enrollment but which cannot be measured directly at the sectoral level. The equations are estimated using weighted least squares; the weights are the population counts (black male teens, ages 16-19). Data are drawn from the state volumes of the 1950 census.

The regression results confirm our hypotheses: the coefficient of ELEM is negative in the enrollment regression and positive in the labor force participation regression. Every ten percentage point decrease in the proportion of teens enrolled in the elementary grades increases the enrollment rate (holding the sector and state constant) by 2.8 percentage points and decreases the labor force participation rate of in-school teens by 6.6 percentage points.

We use the regression coefficients to calculate counterfactual enrollment and participation rates based on different assumptions about the value of ELEM. In 1950, the average value of ELEM for southern black male teens was 0.376. Had ELEM in 1950 been equal to its value in 1970 (8.5 percent) the estimated participation rate of southern black male teens in 1950 would have been 46.1 percent. The difference between the actual and counterfactual 1950 participation rates accounts for 59 percent of the decline in black teen participation between 1950 and 1970.

In 1940, the value of ELEM for southern black male teens was 51.2 percent. Comparable figures for the pre-1940 period are not available, but it is possible to estimate the value of ELEM from information on the age structure of enrollments and the number of pupils enrolled in high school. Using such data we estimate that ELEM was 58.9 percent in 1930. Thus the percent of black teens enrolled in the high school grades had been increasing for at least two decades prior to 1950, producing a decline in labor force participation that predated the mechanization of cotton agriculture and the minimum wage and which could have been expected to continue after 1950.

Racial and regional differences in black teen participation can also be attributed to differences in the value of ELEM. The value of ELEM for southern white male teens was 11.8 percent in 1950. Had there been no racial difference in ELEM in the South in

1950, we estimate that 48 percent of black male teens would have been participating in the labor force, 7 percentage points less than the actual participation rate of southern white male teens in 1950. The fact that black teens were less likely to participate in the labor market than white teens, holding ELEM constant, suggests that racial discrimination may have played a role in limiting labor market opportunities for southern blacks before 1950.

In 1950 the average value of ELEM for nonsouthern black male teens was 15.7 percent. Had the value of ELEM for southern black teens equalled its value outside the region, we predict a participation rate of exactly 50 percent. Since the participation rate of nonsouthern black teens in 1950 was 40.4 percent (compared with 60.8 percent in the South), the regional difference in ELEM explains 52.9 percent [=10.8/20.4] of the regional difference in participation rates.

The decline in labor force participation of southern black male teenagers after 1950 has frequently been attributed to the effects of agricultural mechanization and the minimum wage. The post-1950 decline in participation, however, originated before the diffusion of the mechanical cotton picker and the Fair Labor Standards Act of 1938. A key factor behind the long-term decline was the long-term increase in educational attainment of successive cohorts of southern black children. Once large numbers of black male teens desired to graduate from high school, low rates of school enrollment and high rates of in-school labor force participation were no longer feasible.

Although a significant part of the post-1950 decline in participation was the continuation of a long-term trend, census data also indicate that the decline in participation accelerated after 1950, and that this accelerated decline was reflected primarily in reductions in labor force activity of black teens who had left school. It may be that the post-1950 decline in participation among black teen dropouts was largely due to agricultural mechanization and the minimum wage. Black teen dropouts in the 1950s and 1960s, however, were increasingly undereducated compared with the average young worker. Recent work on the structure of earnings in the United States suggests that the overall demand for educated workers has risen more or less steadily in the post-World War Two period. The corresponding reduction in the demand for less-educated labor, including black teen dropouts, could have produced a decline in their participation, independent of specific effects of agricultural mechanization or the minimum wage.

HEDONIC WAGES AND LABOR MARKET INTEGRATION: THE U.S., 1890-1903

William A. Sundstrom
Santa Clara University

Joshua L. Rosenbloom
University of Kansas

Introduction

In the decades after the Civil War, the completion of national rail and telegraph networks promoted the emergence of broad regional or national markets in many sectors of the economy (James [1978], DuBoff [1982], Chandler [1977]). The persistence of large geographic wage differentials during this era, however, suggests that labor markets may not have participated in the general trend toward increasing national integration (Rosenbloom [1990], Coelho and Shepherd [1976]). An alternative interpretation of these wage differentials is that they reflect compensating variations for geographic differences in other job or locational characteristics. Although historical

studies of U.S. labor markets have generally focussed on hourly or daily wages as the price of labor, wages are only one of the factors that may have influenced the supply of and the demand for labor.

In this paper we extend previous studies of geographic integration in late nineteenth-century labor markets by explicitly considering inter-city variations in both wages and hours of work. Hours of work are just one of the non-wage job characteristics that could have influenced the supply of and demand for labor, but the attention they were accorded by both workers and employers at the time suggests they were an especially important dimension of late nineteenth-century employment relationships. In this era a 60 hour work week (10 hours a day, six days a week) was common, and the movement for shorter hours was a potent rallying cry among both union and nonunion workers (Cahill [1932], Roediger and Foner [1989], Hunnicutt [1988]). Most employers were adamantly opposed to efforts to shorten hours, however, believing that such reductions would substantially increase production costs. Reflecting the importance that both sides attached to the length of the work day, hours of work figured prominently as a source of labor conflict in the late nineteenth century.

Using occupation-specific wage and hours data for 10 occupations in 38 cities from 1890 to 1903, we find no evidence that geographic variations in wages can be explained as a compensating variation for differences in hours of work. In fact, for the skilled and semi-skilled building trades occupations that we consider, hours of work varied in a manner that systematically reinforced rather than offset variations in wage rates. In other words, for these occupations wage differentials alone understate the extent of geographic variation in the rewards to labor. It is possible, of course, that the observed variations in both wages and hours were compensated by variations in other job characteristics or by locational amenities. While data to test this hypothesis are more limited, we argue that the available evidence is not consistent with such a conjecture.

Hours and Wages in Labor Market Equilibrium

The joint determination of hours and wages is a special case of the theory of hedonic prices. Because hedonic equilibrium involves a process of matching and sorting between heterogeneous employers and workers, equilibrium in such a market is characterized not by a single wage-hours pair, but by a range of possible combinations along a market wage-hours schedule. Although the precise shape of the equilibrium hedonic wage-hours schedule cannot be determined, Kinoshita [1987] has shown that under fairly general assumptions it is possible to derive testable restrictions on its shape. Specifically, if workers in one location both earn higher real weekly pay and work fewer hours per week than workers in another location, then the two places cannot be part of a single integrated market, other job and locational characteristics held

equal. Intuitively, workers who earn more money for fewer hours of work are unambiguously better-off, while employers who must pay more for fewer hours of work are unambiguously worse-off. By definition such a situation cannot persist within an integrated market.

Thus in an integrated market weekly earnings and hours should be positively related. Suppose $E = M(h)$ is the market determined hedonic earnings-hours schedule, where E is weekly earnings and h is weekly hours. Then market integration requires $dM/dh > 0$. The restriction can be restated in terms of the wage rate for h hours of work, $W(h)$, by noting that $M(h) = hW(h)$. Thus

$$dM(h)/dh = d\{hW(h)\}/dh = W(h) + h[dW(h)/dh] = W(h)[1 + \Theta] > 0$$

where Θ is the elasticity of $W(h)$ with respect to h . Since $W(h)$ is positive, this inequality is equivalent to the restriction $\Theta > -1$. Another way to interpret $\Theta > -1$ is to say there is a compensating wage differential for longer hours. If we find that $\Theta < -1$, holding constant all other job characteristics, then the data points cannot all be part of the same market. This restriction on the shape of the equilibrium wage-hours relationship suggests a natural test of market integration making use of data on the joint variation in wages and hours. Given the likelihood of a certain amount of random variation in wages or hours considered separately, finding $\Theta < -1$ is stronger evidence of a lack of integration than simply finding that wages and/or hours vary across locations. On the other hand, $\Theta > -1$ is necessary but not sufficient to prove that the market is in fact integrated.

Geographic Variation in Wages and Hours

Shortly after the turn of the century the U.S. Commissioner of Labor undertook the collection of annual data on nominal wages and weekly hours of work in a wide range of industries and occupations. The results of this study were reported in the 19th Annual Report of the Commissioner of Labor (U.S. Department of Commerce and Labor [1905]; hereafter called the 19th Annual Report), which published average wages and hours by industry and occupation from 1890 to 1903. Because data were collected by industry and occupation they have the desirable feature of controlling for variations arising from differing working conditions or employment practices across industries or among occupations within an industry.

From the available data we selected a sample of 10 occupation covering several industries and a variety of skill levels. Four occupations each were drawn from the building trades and foundry and machine shop employment. In addition we included contract and municipal street and sewer construction laborers. Five of the occupations may be classified as highly skilled (bricklayers, carpenters, machinists, iron molders, and pattern makers), four as unskilled (building trades laborers, foundry and machine shop laborers, and the two groups of street and sewer workers), and one as semi-skilled (hod carriers). To ensure reasonable comparability in

geographic coverage we restricted our sample to those 38 cities for which data were available for at least 5 of the selected occupations. To obtain real wages that control for locational cost of living, we used Haines's (1989) city and state cost-of-living indexes for 1890.

The extent of inter- and intra-regional variation in real wages and hours of work revealed by the 19th Annual Report data is summarized in Table 1, which shows the unweighted mean and standard deviation of real wages and hours in 1890 for each occupation by region. Considering wages first, it is apparent that there were pronounced interregional variations. The West, with wage levels 10 to 30% above those in the Northeast, was consistently a high wage region. Meanwhile wages were consistently lowest in the Southeast. These interregional differentials were most pronounced in the building trades, but were also apparent for other occupations. Within the South a considerable differential is apparent between cities in the South Central and Southeast regions. Within the northern part of the country, average wages in most occupations were slightly higher in the North Central region than they were in the Northeast. Despite the near equalization of wages across these two regions, the large size of the within-region standard deviation indicates that the North was still characterized by a considerable degree of intercity wage variation.

Table 1:
Average Real Wages and Hours by Occupation and Region, 1890
(Within Region Standard Deviations In Parentheses)

	Real Wage (cents per hour)						Weekly Hours					
	NE	NC	SE	SC	W	US	NE	NC	SE	SC	W	US
<u>Building Trades Workers</u>												
Bricklayers	41.4 (6.4)	45.4 (7.0)	26.1 (5.3)	48.5 (11.6)	57.1 (3.3)	42.3 (4.8)	53.6 (2.1)	53.3 (2.8)	59.1 (4.1)	55.6 (2.8)	52.4 (3.7)	54.6 (4.1)
Carpenters	27.3 (4.9)	27.0 (4.6)	19.5 (3.7)	27.9 (2.2)	31.6 (0.1)	26.3 (5.2)	54.6 (3.6)	55.5 (4.9)	59.4 (1.3)	56.7 (2.5)	55.1 (1.6)	55.9 (3.9)
Hod Carriers	22.6 (5.0)	23.4 (5.7)	13.9 (0.0)	20.5 (5.8)	28.9 (2.5)	22.5 (5.7)	53.6 (3.3)	53.4 (4.8)	58.7 (1.8)	56.9 (3.0)	56.3 (5.2)	54.6 (4.1)
Laborers	15.4 (1.1)	16.5 (2.7)	8.2 (1.7)	14.1 (2.7)	17.8 ^a	15.0 (3.3)	58.5 (2.4)	59.2 (1.4)	60.0 (0.0)	56.8 (3.0)	60.0 ^a	58.8 (2.0)
<u>Foundry and Machine Shop Workers</u>												
Machinists	23.7 (2.5)	24.3 (3.4)	23.8 (1.4)	29.5 (2.2)	27.3 (0.9)	25.0 (3.3)	59.8 (0.3)	59.2 (0.9)	60.0 (0.0)	59.7 (0.4)	59.6 (0.4)	59.6 (0.6)
Iron Molders	25.7 (2.4)	26.9 (3.1)	23.6 (2.7)	30.1 (0.7)	30.3 (0.6)	26.9 (3.1)	59.7 (0.5)	59.6 (0.5)	60.0 (0.1)	59.5 (0.8)	58.3 (2.8)	59.5 (1.0)
Pattern Makers	27.6 (2.4)	28.6 (3.4)	24.2 (4.4)	30.3 (2.4)	30.3 (1.3)	28.3 (3.2)	59.5 (1.1)	59.2 (0.8)	59.9 (0.3)	59.5 (0.9)	59.5 (0.4)	59.4 (0.9)
Laborers	14.6 (1.8)	15.4 (2.1)	10.8 ^a	14.4 (3.0)	17.5 (1.3)	15.0 (2.4)	59.7 (0.5)	59.5 (0.7)	59.7 ^a	59.6 (0.3)	59.6 (0.5)	59.6 (0.5)
<u>Street and Sewer Construction Laborers</u>												
Contract	15.0 (3.3)	16.2 (2.3)	15.0 (0.1)	14.5 (4.7)	17.1 ^a	15.5 (2.6)	59.1 (2.2)	57.8 (4.5)	59.0 (1.0)	60.0 (0.0)	60.0 ^a	58.7 (3.0)
Municipal	16.6 (2.8)	17.5 (3.1)	12.2 (2.7)	13.6 (1.5)	21.5 ^a	15.9 (3.4)	57.2 (4.0)	55.5 (6.2)	59.5 (1.0)	58.8 (2.5)	48.0	56.9 (4.8)

^a Data available for only one city for this occupation and region.

Notes: Regional definitions are as follows: NE (Northeast)--Baltimore, Boston, Buffalo, Newark, New York, Philadelphia, Pittsburgh, Providence, Rochester, and Williamsport; NC (North Central)--Chicago, Cincinnati, Cleveland, Detroit, Dubuque, Evansville, Indianapolis, Milwaukee, St. Louis, St. Paul, Sioux Falls; SE (Southeast)--Atlanta, Augusta, Charleston, Jacksonville, Richmond, Wilmington; SC (South Central)--Little Rock, New Orleans, Louisville, Memphis, Mobile, Nashville; W (West)--Portland, San Francisco, Seattle.

Sources: U.S. Department of Commerce and Labor (1905); Haines (1989).

Turning to hours of work, there was remarkable uniformity in the work week for skilled and unskilled foundry and machine shop workers. A work week of nearly 60 hours was typical in every region of the country, and there was very little within-region variation around this average. In sharp contrast, hours of work were both shorter and more variable among skilled and semi-skilled building trades workers and municipal street and sewer workers. The situation for building trades laborers and contract street and sewer workers lies somewhere between these two extremes--on average the work week in each region remained close to 60 hours, but there was somewhat more within region variability. While there was a downward trend in hours of work after 1890, the intra- and interregional patterns of hours variation apparent at this date remained largely unchanged in later years.

Market Integration in the Late Nineteenth Century

As we have seen, the appropriate test of market integration when both wage and hours vary hinges on whether the wage-hours elasticity along the market wage-hours schedule is less than -1. To determine this elasticity we estimated the following relationship separately for each occupation:

$$\ln(W_i/P_i) = \alpha + \Theta \ln(H_i) + \epsilon_i, \quad (1)$$

where i indexes cities, W is the nominal wage, P is the cost of living, H is weekly hours, and ϵ is a stochastic error term assumed to have mean equal to zero. In this equation Θ is the elasticity of real wages with respect to hours of work along the hedonic wage-hours schedule and should be greater than -1 if the market is integrated.

The first two columns of Table 2 report the ordinary least squares estimates of Θ for each occupation in 1890. We focus here on 1890 because it is the only year for which we have city cost-of-living data, but we consider other years in the analysis below. Overall the results are consistent with the impression based on interregional wage differentials alone that a unified national labor market did not exist in 1890. For the skilled and semi-skilled building trades workers and for municipal street and sewer workers hours variations reinforce wage variations, rather than offsetting them. In each of these instances the estimated elasticity is statistically significantly less than -1, indicating that weekly earnings were greatest precisely in those places where hours were shortest. Such a pattern cannot have been drawn from points along the equilibrium hedonic wage-hours schedule of a single unified market. For all other occupations, the regression estimates indicate that there was in effect no systematic relationship between intercity variations in wages and hours. One cannot reject the null hypothesis that there was no compensating differential for longer hours.

In addition to examining integration at the national level, the joint variation of wages and hours can be used to examine more precisely intraregional integration by estimating equation (1) separately for regions and subregions. The results of such regressions, not reported here, show some regional differences, but no tendency in any region or occupation toward significant compensating differentials for longer hours in any occupation. In the South, the hypothesis of market integration is decisively rejected for the skilled and semi-skilled building trades workers.

Table 2:
Estimated Wage-Hours Elasticities (Θ) in 1890

	Separate occupational regressions		Regressions with city fixed effects	
	Elasticity	Standard error	Elasticity	Standard error
<u>Building Trades Workers</u>				
Bricklayers	-2.256 ^a	0.528	-1.631 ^a	0.345
Carpenters	-2.024 ^a	0.368	-1.338	0.347
Hod Carriers	-2.090 ^a	0.528	-2.017 ^a	0.354
Laborers	-0.769	1.763	-0.278	0.924
<u>Foundry and Machine Shop Workers</u>				
Machinists	0.762	2.195	2.121	2.278
Iron Molders	-1.157	1.328	0.310	1.543
Pattern Makers	-1.113	1.493	-1.382	1.723
Laborers	1.544	3.975	1.472	3.276
<u>Street and Sewer Construction Laborers</u>				
Contract	-0.684	0.695	-0.865	0.550
Municipal	-1.817 ^a	0.360	-1.245	0.326

^a Significantly less than -1 at the 95% confidence level

Notes: The wage-hours elasticities in the third column are derived from a single regression that includes separate intercepts for each city and occupation.

Source: U.S. Department of Commerce and Labor (1905); Haines (1989).

In the North, the estimates are less precise, but the point estimates also reject market integration for these occupations, strongly suggesting that intercity variations in labor market conditions were greater for skilled and semi-skilled building trades workers than for other occupations.

Wages, Hours, and Other Job Characteristics

Our focus on wages and hours in hedonic equilibrium is strictly appropriate only if other dimensions of the employment relationship and locational amenities are held constant across cities. In addition to hours of work, however, jobs may be differentiated by such factors as the pleasantness of working conditions, skill requirements, and the risks of injury, death, or unemployment. While variations in these factors are undoubtedly important across industries and occupations, it is less plausible that variations in job characteristics were important within the narrow industry and occupation categories that we consider in this paper. Workers may also be concerned with more general locational characteristics, such as city size, climate, or other factors affecting the quality of life. While it is not possible to adequately quantify all such factors, it is possible to get some idea of their potential significance by pooling data for different occupations and controlling for fixed city effects. Since location-specific amenities probably affect all workers in a city similarly, they should have a systematic effect on wages across occupations in a city. Assuming that this effect is proportional to wages, we can control for location-specific amenities by pooling the data for all ten occupations in our sample and estimating the following relationship:

$$\ln(W_{ij}/P_i) = \alpha_j + \beta_i + \Theta_j \ln(H_{ij}) + \epsilon_{ij}, \quad (2)$$

where i indexes cities, j indexes occupations, and the coefficient β is an estimate of the magnitude of location-specific amenities.

The third and fourth columns of Table 2 report the wage-hours elasticities and their standard errors obtained from ordinary least squares estimates of equation (2). Comparing these elasticities to those in the first two columns of the table suggests that there is little qualitative change in the results. The wage-hours elasticities are increased for most occupations and the precision of the estimated parameters declines, but the same occupational pattern of differences in the wage-hour relationship is evident. For the skilled and semi-skilled building trades and the municipal street and sewer workers, nationally and by region, Θ is almost always less than -1, while for the other occupations it is generally greater than this value. Unless locational amenities are hypothesized to affect some occupations very differently from others, it is thus difficult to imagine that they could provide a means of explaining intercity variations in wages and hours.

Similar results obtain when this analysis is extended to regional subsamples.

Market Integration in the Long Run

So far we have concentrated on wage and hours variation in a single year--1890--because this is the only year for which relative price level estimates are available. It is possible, however, that the pattern of intercity variation in this year, and especially the occupational differences in the wage-hours relationship, reflect a temporary disequilibrium caused by localized supply or demand shocks. To address this possibility, we looked at how the patterns of wage and hours variations evolved in the years after 1890, estimating equation (2) for other years. The results, not reported here, exhibit a general persistence through time: where a systematic wage-hours relationship existed in 1890 it persisted largely unchanged. Of course, the stability of the estimated wage-hours relationship across cities need not imply that relative real wages and hours of work remained the same in individual cities. In fact, however, the geographical pattern of real wage and hours variations was remarkably stable. One can show that the deviation of real wages from their cross-city mean at the beginning of the period is a very good predictor of the deviation of real wages at the end of the period. This persistence is somewhat weaker for hours in some occupations, but was again quite strong for the building trades.

Conclusion

Throughout the late nineteenth and early twentieth centuries, workers sought both to raise wages and shorten their hours of work. In this paper we have attempted to incorporate information about both of these variables into an assessment of the extent of labor market integration from 1890 to 1903. In none of the occupations that we considered did we find evidence of systematic compensating differentials associated with differences in the length of the work week. Thus in no case does incorporating variation in hours of work reverse conclusions about integration based solely on the magnitude of wage differentials. On the other hand, combining data on wage and hours variations sharpens the impression that there were significant occupational differences in the extent of geographic integration. In the markets for carpenters, bricklayers, hod carriers, and municipal street and sewer workers not only were wages more variable across cities, but hours were shorter in those places where weekly earnings were higher, meaning that differences in wages alone understate differences in the rewards to labor.

(Full-length version of paper and references available from authors upon request.)