50TH CELEBRATORY VOLUME

Edited by Solomon W. Polachek
and Konstantinos Tatsiramos

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

RESEARCH IN LABOR
ECONOMICS

VOLUME 50

50TH CELEBRATORY VOLUME

RESEARCH IN LABOR ECONOMICS

Series Editor: Solomon W. Polachek

IZA Co-Editor: Konstantinos Tatsiramos

Recent Volumes:

Volume 34:	Informal Employment in Emerging and Transition Economies *Edited by Hartmut Lehmann and Konstantinos Tatsiramos*
Volume 35:	35th Anniversary Retrospective *Edited by Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 36:	Research in Labor Economics *Edited by Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 37:	Labor Market Issues in China *Edited by Corrado Giulietti, Konstantinos Tatsiramos and Klaus F. Zimmermann*
Volume 38:	New Analyses in Worker Well-Being *Edited by Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 39:	Safety Nets and Benefit Dependence *Edited by Stephane Carcillo, Herwig Immervoll, Stephen P. Jenkins, Sebastian Konigs and Konstantinos Tatsiramos*
Volume 40:	Factors Affecting Worker Well-Being: The Impact of Change in the Labor Market *Edited by Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 41:	Gender Convergence in the Labor Market *Edited by Solomon W. Polachek, Konstantinos Tatsiramos and Klaus F. Zimmermann*
Volume 42:	Gender and the Labor Market *Edited by Solomon W. Polachek, Konstantinos Tatsiramos and Klaus F. Zimmermann*
Volume 43:	Inequality: Causes and Consequences *Edited by Lorenzo Cappellari, Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 44:	Income Inequality Around the World *Edited by Lorenzo Cappellari, Solomon W. Polachek and Konstantinos Tatsiramos*
Volume 45:	Skill Mismatch in Labor Markets *Edited by Solomon W. Polachek, Konstantinos Pouliakas, Giovanni Russo and Konstantinos Tatsiramos*

Volume 46: Transitions Through the Labor Market: Work, Occupation, Earnings and Retirement
Edited by Solomon W. Polachek and Konstantinos Tatsiramos

Volume 47: Health and Labor Markets
Edited by Solomon W. Polachek and Konstantinos Tatsiramos

Volume 48: Change at Home, in the Labor Market, and on the Job
Edited by Solomon W. Polachek and Konstantinos Tatsiramos

Volume 49: Workplace Productivity and Management Practices
Edited by Solomon W. Polachek and Konstantinos Tatsiramos

This page intentionally left blank

EDITORIAL ADVISORY BOARD

This page intentionally left blank

RESEARCH IN LABOR ECONOMICS VOLUME 50

50TH CELEBRATORY VOLUME

EDITED BY

SOLOMON W. POLACHEK
State University of New York at Binghamton, USA

And

KONSTANTINOS TATSIRAMOS
University of Luxembourg and Luxembourg Institute of Socio-Economic Research (LISER), Luxembourg

emerald PUBLISHING

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2023

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-80455-126-4 (Print)
ISBN: 978-1-80455-125-7 (Online)
ISBN: 978-1-80455-127-1 (Epub)

ISSN: 0147-9121 (Series)

CONTENTS

LIST OF CONTRIBUTORS

Joseph G. Altonji	Yale University, USA
George J. Borjas	Harvard University, USA
Hugh Cassidy	Kansas State University, USA
Nicole M. Fortin	University of British Columbia, Canada
Pietro Garibaldi	University of Torino, Italy
Daniel S. Hamermesh	University of Texas, USA
Grant Hayes	University of Chicago, USA
John Eric Humphries	Yale University, USA
James Jedras	Boston University, USA
Thomas J. Kniesner	Claremont Graduate University, USA
Edward P. Lazear	Deceased, Stanford University, USA
Thomas Lemieux	University of British Columbia, Canada
Neil Lloyd	University of Warwick, UK
Shelly Lundberg	University of California Santa Barbara, USA
Espen R. Moen	Norwegian Business School, Norway
Michał Myck	Centre for Economic Analysis – CenEA, Poland
David Neumark	University of California Irvine, USA
Monika Oczkowska	Centre for Economic Analysis – CenEA, Poland
Christopher A. Pissarides	London School of Economics, UK
Kathryn Shaw	Stanford University, USA
Robert Topel	University of Chicago, USA
Giannina Vaccaro	University of Lausanne, Switzerland
W. Kip Viscusi	Vanderbilt University, USA
Ling Zhong	Cheung Kong Graduate School of Business, China

This page intentionally left blank

FOREWORD

RESEARCH IN LABOR ECONOMICS: THE EARLY YEARS AND NOW

Around 1974, while I was on the faculty at the University of Massachusetts, I received a telephone call (remember email was still years away) from Paul Uselding. Paul, who had been a graduate school classmate of mine at Northwestern, was then an assistant professor of economics at the University of Illinois – Urbana-Champaign. Paul was also the editor of an annual monograph series, *Research in Economic History*, which was published by JAI Press. His series published economic history papers that were longer than those published by conventional economic history journals.

Paul told me that he and the publisher of JAI Press, Herbert Johnson, had decided that his series was sufficiently successful that it could serve as a prototype for similar series in different fields of economics. Would I be interested in serving as editor for an annual series in labor economics? Having received my PhD only four years earlier and still trying to build a record of accomplishments for myself, I jumped at the idea and *Research in Labor Economics* (RLE) was born.

As I said in the preface to the first volume, RLE "will provide a forum for original contributions whose specific subject matters and methodological approaches will be governed only by composition of the materials submitted by the profession to me for possible inclusion." My goal was to attract papers whose level of treatment would be comparable to (or exceed) those found in leading economics journals. I hoped that three types of papers would be published in RLE: results of completed or ongoing research, critical survey articles on important topics, and symposia consisting of several pieces of research on related topics.

Our first volume, published in 1977, set a baseline standard of excellence. Volume 1 contained papers by luminaries including Sherwin Rosen, one of the leading labor theorists of his generation, Frank Brechling, one of my dissertation advisors, and then young leading empirical labor economists John Pencavel (Stanford) and Richard Freeman (Harvard).

By 1992, we had published 15 more volumes of RLE including three that contained symposia on "Labor Economics and Public Policy," "New Approaches to Labor Unions," and "Evaluating Manpower Training Programs."[1] I was fortunate enough to attract high quality papers by leading labor economists from

[1]Within a few years the profession realized the sexism of the term "Manpower Training Programs" and this term was replaced by the term "Employment and Training Programs."

different generations. At the risk of alienating authors whose names I have omitted, the table below illustrates several of the notable economists who contributed one or more papers to RLE during the period. If one peruses this table, one will find several Nobel Prize winners, winners of the Jacob Mincer award for Lifetime Contributions to Labor Economics presented by the Society of Labor Economics, and winners of the IZA prize for labor economics, as well as editors of the *American Economic Review*. I am particularly proud that two of the pioneers of empirical labor economics research, H. Gregg Lewis and Jacob Mincer, were contributors to these volumes, as well as the distinguished econometrician Zvi Griliches.

Selected Research in Labor Economics Authors: Ehrenberg Era 1978–1992*.

Joe Altonji (Yale) – I	Lawrence Kahn (Cornell)
Joshua Angrist (MIT) – N	Ed Lazear (Chicago) – I, M
Orley Ashenfelter (Princeton) – A, I, M	H. Gregg Lewis (Chicago)
Rebecca Blank (Wisconsin)	James Medoff (Harvard)
Francine Blau (Cornell) – I, M	Jacob Mincer (Columbia) – I, M
George Borjas (Harvard) – I	Olivia Mitchell (Pennsylvania)
Charles Brown (Michigan)	Robert Moffitt (Johns Hopkins) – A
Hank Farber (Princeton) – M	Dale Mortensen (Northwestern) – I, M, N
Gary Fields (Cornell) – I	Ronald Oaxaca (Arizona)
Richard Freeman (Harvard) – I, M	John Pencavel (Stanford) – M
Zvi Griliches (Harvard)	Sherwin Rosen (Chicago)
Daniel Hamermesh (Michigan State) – I, M	Burton Weisbrod (Northwestern)
James Heckman (Chicago) – M, N	Finis Welch (Texas A&M)-M
George Johnson (Michigan)	

*Affiliation is the institution for which the individual is best known.A – editor of the AER, I – winner of the IZA Prize in Labor Economics, M – winner of the SOLE Jacob Mincer Award, N – winner of the Nobel Prize in Economics.

By the early 1990s, editing RLE had become more of a chore for me than an honor. My responsibilities at Cornell and the number of my own PhD students that I was supervising had increased. My research interests were shifting away from labor economics to the economics of higher education. As economists became more preoccupied with publishing in "top 5" journals I was finding it more difficult to attract papers to RLE. Conventional labor economics journals were expanding the size of papers they published and papers that in previous years would have gone to RLE were now going to these journals. We had also reached a time in the profession when the numbers and quality of economists publishing labor economics research in foreign countries was increasing and I had a sense that a new editor, who was more attuned to the changes occurring, would help maintain the viability of RLE and perhaps take it to an even higher level.

My friend Sol Polachek, who by then was a professor at Binghamton University, which is only an hour away from Cornell and is my alma mater, was an obvious choice to replace me. By doing so I would kill two birds with one stone; get a high-quality scholar to replace me and take RLE to new heights and give some publicity and name recognition to my undergraduate institution.

This 50th volume of RLE shows how successful Sol and his coeditors have been. Over the years they have arranged for articles in RLE to be indexed in several indices from which Google Scholar draws. As is well known, Google Scholar is a source that economists regularly use to aid them make judgment about the quality of a scholar at hiring, tenure, and promotion times. My failure to pursue this strategy had limited the flow of manuscripts from economists who were concerned that an article in RLE was not viewed as important as an article in a conventional journal. Sol arranged for a new publisher for the series and the new publisher, Emerald Press, helped RLE to become indexed in the Thomson Reuters Book Citation Index.

By involving IZA in the sponsorship of RLE and enlisting foreign coeditors who are members of IZA, Sol has successfully increased the number of articles RLE has published by labor economists from around the world in several ways. First, approximately 1600 labor economists who are affiliated with IZA now receive notices of the publication of each volume of RLE. The knowledge of what RLE is all about undoubtedly encouraged some recipients to purchase issues of RLE and others to submit papers to RLE. Second, the appointment of European coeditors, connected with IZA, has provided help for Sol in soliciting papers from younger foreign scholars who he may be less aware about. Third, IZA has provided funding for periodic conferences that lead to conference volumes.

The first coeditor was Oliver Bargain (now at the Bordeaux University). He produced a conference volume on microsimulation models, including a forward by Tony Atkinson. That volume was very well received. In 2007 Konstantinos Tatsiramos (Professor of Labor Economics jointly at the University of Luxembourg and the Luxembourg Institute of Socio-Economic Research (LISER)) took over as the coeditor and continues to the present. Sol reported to me that Konstantinos has been truly instrumental in raising RLE's image and stature. Not only has he assiduously obtained conference funding from IZA and external sources, but his contacts with European scholars increased both the quality of the submissions to RLE and the quality of the referees used by RLE. In addition, his eye for detail has enhanced the quality of the final accepted papers.

Among the conference volumes were ones on Immigration, Ethnicity, Child Labor, Informal Employment, Gender Inequality, Skill Mismatch, Health, and Workplace Practices. In addition, RLE's 35th anniversary volume republished the 20 most cited RLE articles that appeared in earlier volumes, that included a new preface by each author containing anecdotes about the paper's original development and subsequent impact.

RLE has continued to publish papers from well-known senior economists and rising younger economists; the latter much more frequently than when I was editor, and they have come from economists not affiliated with US institutions more frequently than when I was editor. Again, at the risk of alienating people who I did not include on the list, the next table displays some well-known economists who published in RLE during the Polachek years. On this list, are several winners of either the Mincer award or the IZA award, both for lifetime achievements, winners of the John Bates Clark award presented to the "best economist under the age of 40" by the American Economic Association, and two authors who have gone on to win the Nobel Prize in Economics.

Selected Research in Labor Economics Authors: Polachek Era.*

Daron Acemoglu (MIT) – J, E	Chinhui Juhn (Houston)
Christian Belzil (Ecole Polytechnique)	Larry Katz (Harvard) – E, I
Dan Black (Chicago)	Alan Krueger (Princeton) – I
John Bound (Michigan)	Thomas Lemieux (British Columbia)
David Card (Berkeley) – I, J, M, N	Shelly Lundberg (UCSB) – I
Andrew Clark (Paris School of Economics)	Richard Murnane (Harvard)
Janet Currie (Princeton)	Paul Oyer (Stanford) – E
William Darrity (Duke)	Daniele Paserman (Boston U)
Greg Duncan (UC Irvine)	Steve Pischke (LSE)
Joseph Ferrie (Northwestern)	Robert Topel (Chicago) – E
David Figlio (Northwestern)	Jane Waldfogel (Columbia)
Robert Haveman (Wisconsin)	Bruce Weinberg (Ohio State)
Barry Hirsch (Georgia State)	Yoram Weiss (Tel Aviv) – M, E

*Affiliation is for the most recent institution at which the individual was located. I – winner of the IZA Prize in Labor Economics, M – Winner of the SOLE Jacob Mincer Award, J – Winner of the John Bates Clark Medal, and E – Editor of either the QJE, JPE Econometrica, or JOLE, N – winner of the Nobel Prize in Economics.

In conclusion, it is with great pleasure that I have watched over the years how RLE has grown in stature under the leadership of Sol and his coeditors, as well as all the actions they have taken to keep RLE on such a positive trajectory.

Ronald G. Ehrenberg
August 2021

Ronald G. Ehrenberg is the Irving M Ives Professor Emeritus of Industrial and Labor Relations and Economics and a Stephen H Weiss Presidential Fellow at Cornell University, as well as the Founding Director of the Cornell Higher Education Research Institute (CHERI) from 1999 to 2021.

PREFACE

This marks the 50th volume of Research in Labor Economics. The volume contains 10 original articles each written by stellar senior scholars in labor economics. One is a Nobel Laureate. Each article deals with an aspect of worker well-being, a hallmark subject of concern especially to labor economists. Of these, five deal directly with human capital and potential earnings, four with institutional impediments to time allocation including work, and one important article shows how economics-based search and matching theory should be applied elsewhere, particularly in epidemiological modeling with regard to pandemics, such as the recent COVID-19 pandemic.

Search and matching models are fundamental to economics. Based on search's costs and benefits, these models delineate how individuals sort into a variety of activities including marriage, jobs, schools, fields of study, and numerous other pursuits. More broadly, search and matching models explain unemployment and unemployment duration which have important macroeconomic implications. Independent of economists, epidemiologists use matching to model the propagation of disease. Here individuals meet others, but now potentially become infected with illness, rather than obtaining an economic benefit. In the first article, according to Pietro Garibaldi, Espen Moen, and Christopher Pissarides, these latter epidemiological models in many ways have similarities to search and matching models, but are distinctly different mostly because they neglect individual cost and benefit incentives. As an example, rising infection rates make social interactions more expensive inducing individuals to shift away from interpersonal interactions, what we now call social distancing. As a result, meeting probabilities are reduced. Incorporating these costs affect equilibrium disease evolution, and hence appropriate national policies to combat disease spread. As such, there is much epidemiologists can learn from economists.

It is well known that human capital is the backbone of economic success. Countries with a more educated populace are richer and individuals with more schooling earn more. Yet it is difficult to measure the benefits of school because of endogeneity issues related to sorting. If smarter students stay in school longer, for example because they benefit more than the less able, simply computed rates of return are overstated. And, of course, the opposite is true if the more able have higher returns in the labor market than in the classroom. However, other sorting issues arise if individuals vary in their types of ability. In this case heterogeneity in learning can manifest itself both in the levels and types of education a person receives. If so, how individuals sort into fields of study is in part dependent on the distribution of innate attributes including ability. As a result, there are several relevant technical econometric issues that need to be addressed to take account of the resulting endogeneity. In the next article, Joseph Altonji, John Humphries, and Ling Zhong address these issues and estimate returns to advanced postgraduate degrees for men and women. To do this they adopt what they call a "fixed-effects combination of group (FEcg)" estimator that gets at endogeneity related to an individual's choice of graduate study based on ability and preferences. They make four contributions. First, they estimate earnings premia by gender associated with 19 advanced degrees. Second, they decompose these earnings premia into portions based on wages and hours. Third, they compute internal rates of return. Finally, they examine job satisfaction.

Higher wages resulting from human capital acquisition, particularly from advanced postgraduate schooling, is an example of compensating wage differentials, as it reflects recompense for time and money spent investing. But other job attributes including both amenities and unpleasant characteristics can affect wages, as well. To get at these, one can apply hedonic estimation models in a Mincer-type earnings equation. Doing so, paying particular attention to the risk of death or injury on the job, enables one to back out the value of a statistical life (VSL), the value of a statistical injury (VSI), and the value of a statistical life year (VSLY). In the next article, Thomas Kniesner and W. Kip Viscusi describe this process and then survey existing studies. They compare the value of life estimates in the United States to other countries and they compare the value of life estimates between men and women. They also examine other possible applications and show how the approach can be used to evaluate welfare policies. Finally, they apply the approach to evaluate the cost of COVID-19 in the United States and worldwide.

Whereas some wage variation arises because of compensating wage differentials, some variation can also come about because of market imperfections. One such imperfection is discrimination, whereby equally competent and equally productive workers receive disparate wages despite desiring exactly the same job. This is often stated to be the case by race, as (full-time) black men earn approximately 78 cents on the dollar compared to (full-time) white men, and (full-time) females now earn about 83 cents on the dollar compared to (full-time) males. However, how much of these wage differences are due to compensating differences, and how much due to discrimination, is tricky to ascertain. Men and women earn roughly equal wages upon entering the labor force at the beginning of their respective work lives, but wages diverge over the life cycle as men and women move through their careers. In the next article, David Neumark and Giannina Vaccaro note that some view this as prima facie evidence against gender discrimination, otherwise why would men and women be paid equally well when they first get employed. To counter this argument, they point out a fallacy to this assertion. Given that observed earnings net out human capital investments (i.e., observed earnings equal potential earnings minus the cost of on-the-job human capital investments), they argue that women should be paid more than men because women's on average lower lifetime work leads them to obtain less human capital early on. Thus, in the absence of discrimination, women should be earning more than men at career onset. Thus, they claim initial wage parity can still imply discrimination.

Clearly there still remain a number of other ambiguous and unanswered questions regarding gender differences. In the next chapter, Shelly Lundberg argues that much of the gender discrimination literature decomposes gender gaps into "discrimination" and "choice." However, choice is difficult to define since what many believe to be choice can really result from discrimination, and what many deem to be discrimination may actually be choice from a perspective "that views the default agent as male" if men and women are motivated differently. Further gender is difficult to define especially now that we realize the very many gender identities psychologists and others study. As such, the article argues for more complex modeling.

Whereas inequality between men and women narrowed, and more quickly in the last 40 years, inequality between the rich and poor has risen dramatically, at least in a number of the most developed countries during the same time period. Interestingly, a 2020 UN Report found inequality (based on the Gini coefficient) rising and falling in roughly equal proportions out of 119 countries between 1990 and 2016. But more developed countries constituted the preponderance of countries where earnings inequality rose. Two-thirds of these (26 out of 39) experienced a widening earnings disparity between top and bottom earners. From the mid-1980s until 2011, the Gini coefficient increased in 16 out of the 21 OECD countries for which long time series were available. Why these increases prevail in developed countries is yet an unanswered question. In the next article, Edward P. Lazear, Kathryn Shaw, Grant Hayes, and James Jedras explore why. Predicated on their unique education-specific industry-based index, they argue in favor of skilled-biased technical change probably stemming from the AI revolution by which productivity and wages rose more for the well educated than the less educated.

Technical change is not the only factor that affects labor market outcomes. Governments do so directly via laws explicitly regulating the market. In the next article, Robert Topel deals with potential legislative overreach. Here bureaucrats stretch their agency's regulatory powers beyond government's original intent. Examples cited include the Equal Employment Opportunity Commission's use of criminal background checks in hiring decisions. Part of this may occur because of imprecise statutory language (Congress in the case of the United States). Often governments do not define precisely the social problems in all its detail and complexity. In Topel's words, part of the problem may be that "bureaucrats attracted to work in the agency ... have a disproportionate belief in its mission." Even so, there is at least one other important factor enabling the propagation of fervorous bureaucrats. Topel shows this can come about if "the evaluation of agency recommendations and punishment of the agency for poor performance are in different hands." One reason is that costs and benefits differ between the two. In any event, overregulation is greatest "when the social problem is least harmful" and "the oversight agency" is weakest.

Laws still matter independent of whether government bureaucrats over- or underregulate. Right to Work (RTW) legislation is one set of laws that potentially have powerful implications for labor. As is well known, RTW states cannot compel workers to join a union and pay dues even if represented under a collective bargaining agreement. Such laws were prevalent in the South, but of late have been passed in such states as Oklahoma, Indiana, Michigan, Wisconsin, West Virginia, and Kentucky. Clearly such legislation diminishes the power of unions, and thus has implications for unionization rates and worker wages. Given their staggered implementation across states and their varied effect across industries, Nicole Fortin, Thomas Lemieux, and Neil Lloyd adopt several identification strategies to measure their effect on unionization and wages in the next article. They find RTW laws lower wages and unionization rates. Further, they find unions increase wages about 35% when using RTW as an instrument in a wage equation.

Regulation isn't the only aspect that affects human behavior. The last two years witnessed one of the biggest unanticipated shocks felt world-wide in over a century. It is now well known that the COVID-19 pandemic dramatically reduced employment. In April 2020, just after the pandemic hit, the United States lost over 20 million jobs. Unemployment soared from 3.5% to 14.7%. Also well known, women were more dramatically affected than men. Their unemployment shot up from 3.6% in January 2020 to 16.1% in April 2020. Less well known is how immigrants, especially undocumented immigrants, fared during the pandemic. In the next article, George Borjas and Hugh Cassidy show a much more precipitous drop in employment for immigrants, but at the same time, a far faster recovery, especially for the undocumented immigrants. Interestingly, Borjas and Cassidy find the type of immigrants' jobs are most responsible for their bigger job losses, as these jobs had O*NET characteristics less inducive to working at home. But even more interestingly, undocumented workers rebounded more quickly than either natives or other immigrants, mostly because they did not qualify for generous pandemic instigated unemployment insurance benefits.

Costs and benefits are affected by world-wide shocks, but personal trauma can do the same, at least at the household level. In the final article, Daniel S. Hamermesh, Michał Myck, and Monika Oczkowska examine the effect of a specific, in some cases expected, and in some cases unexpected, familial shock, namely the death of a husband on a widow's time allocation. Given higher women's life expectancy, widows constitute an important yet understudied demographic group. Though there are slight variations across countries, roughly 50% of women over 70 were widowed compared to about 15% for similarly aged men. Using high-quality time-diary data, the authors analyze time usage for widows in a number of European countries and in the United States. They find newly widowed women spend less time in home production and personal care than married women, but more time sleeping and in other leisure activities. Longer-term widowed women spend even less time in home production but more time watching television.

The success of Research in Labor Economics depends on its authors and the quality of their articles. As this volume attests all authors have stellar reputations and have written significant state of the art articles with important policy implications. For this volume we especially thank Daniel Hamermesh who supported us in the initial stages of its planning. We also congratulate Simon

Jäger on becoming the new IZA Director. He is a noteworthy labor economist whom we are eager to work with as we produce future RLE volumes in conjunction with IZA. In this regard, readers who have prepared manuscripts that meet RLE's stringent standards are encouraged to submit them via the IZA website (http://rle.iza.org).

Solomon W. Polachek
Konstantinos Tatsiramos
Volume Editors

This page intentionally left blank

THE SAM APPROACH TO EPIDEMIC MODELS

Pietro Garibaldi, Espen R. Moen
and Christopher A. Pissarides

ABSTRACT

We discuss the connections between epidemiology models and the search and matching (SAM) approach and draw conclusions about modeling the trade-offs between lockdowns and disease spread. We review the pre-COVID epidemics literature, which was mainly by epidemiologists, and the post-COVID surge in economics papers that use meeting technologies to model the trade-offs. We argue that modeling the decentralized equilibrium with economic trade-offs gives rise to substantially different results from the earlier epidemics literature, but policy action is still welfare-improving because of several externalities.

Keywords: SIR models; matching models; COVID-19; social distancing; herd immunity; externalities

JEL codes: A12; I10; J18; D61-D62

1. INTRODUCTION

Models of infectious diseases share a key feature with a large number of economic models. This is that the utility (positive or negative) that an agent derives from their activities requires contact with another agent in "social space." In economics, this feature of exchange is stronger in frictional markets, in which participants need to search, find, and inspect alternatives before deciding to buy or sell, than in market-clearing neoclassical markets, in which goods are homogenous. A typical example is the labor market, in which participants need to search over alternatives before agreeing to a match. Another is the housing market, in which buyers need to search and inspect houses before buying.

50th Celebratory Volume
Research in Labor Economics, Volume 50, 1–23

ISSN: 0147-9121/doi:10.1108/S0147-912120230000050001

A useful way to formalize these interactions is as a dynamic process during which the individual changes state, e.g., from unemployed to employed, or from a renter to a homeowner.[1] The transition from one state to another can then be modeled as in the "search and matching," or SAM, approach. This approach has been successfully applied initially to models of the labor market and subsequently to several others in which there is need to gather idiosyncratic information before exchange.[2]

The dynamics of epidemics have been modeled as transitions between states long before economists developed this particular methodology (Kermack & McKendrick, 1927). But the use of models of individual behavior driven by incentives, which motivated economists, was not applied to epidemiological models until much later. Our motivation for writing this paper is that we believe that epidemiological models in the spirit of Kermack and McKendrick (1927), have a lot to gain from the insights of the SAM approach, as developed by economists.[3] We critically survey models of infectious diseases that use matching processes to derive the transitions from a healthy (or "susceptible") state to an infectious one. As we will argue, variants of this approach have been used sparingly in the pre-COVID research but exploded during the COVID-19 pandemic.

In economic models the typical SAM situation is one in which contact is the result of some costly activity, such as search, and yields positive returns if successful, or nothing. In contrast to this typical situation, in an epidemic there is a positive probability that contact between two people would involve some cost: the transmission of the disease from an infected to a healthy individual. Putting the economic and epidemiological models together, a trade-off is created. The more activities you pursue in social space to increase your economic payoffs, the higher the risk of infection. Viewed from the lens of the economist, an epidemic introduces a cost to the economic interaction in social space, which shifts economic activities in favor of the ones that do not involve social interaction. For example, shopping activities shift from browsing in shops to buying online. Viewed from the lens of epidemiology, restricting social contact contains the disease but it puts a cost on society, the reduction in utility that could be derived from social interaction.

Before the outbreak of the COVID-19 pandemic in 2019, this trade-off led to a modest literature, mostly written by mathematical biologists, although there are also some papers by economists. Epidemiologists focused mainly on influenza outbreaks, whereas the best-known papers by economists are about HIV transmission. Following the outbreak of COVID-19, however, a very large economics literature emerged, covering all aspects of the pandemic and its economic costs. Our survey covers only a subfield of these literature strands, the one that models transitions between states by making use of a contact function, which could be a matching function as in SAM theory, or any other mathematical representation of contacts.

We summarize the way that this function has been modeled by epidemiologists and economists and show that its form influences the nature of equilibrium and its welfare properties. We argue that although in the labor literature a linear-homogeneous function has received a lot of empirical support, in epidemiology it is more common to use a function with increasing returns to scale (usually the quadratic).[4] This distinction is plausible because in economics contacts are usually used for the exchange of a single good, e.g., a house or a service that only one person can enjoy, whereas in epidemiology a person carrying the virus can pass it on to a large number of contacts. This distinction turns out to be important because an action such as social distancing is much more effective in reducing contacts if the contact function is quadratic than if it is linear. So the policy recommendation due to the quadratic, or any other increasing returns function, is usually for more social distancing, but the exact recommendation depends also on other features of the model, such as heterogeneity or forward-looking behavior.[5]

We split the pre-COVID economics research into two main strands. First, we review the HIV research, working mainly with an SI (susceptible-infected) model. Second, we review models that use Kermack and McKendrick's (1927) susceptible-infected-removed (SIR) model. Key contributions to the HIV research on the SI model are Geoffard and Philipson (1996) and Kremer (1996). These models do not introduce explicitly an aggregate contact function that depends on individual incentives. Yet, they show that the flow from susceptible to infected individuals depends on the number of people currently infected and the prevailing sexual activity, which depends on the HIV infection risk. The SIR model pre-COVID was applied mainly to the modeling of influenza epidemics, and the question usually addressed was how much social distancing should the government legislate during an epidemic. Early models that introduced some kind of contact function in influenza models include Reluga (2010) and Chen, Miahoua, Rabidoux, and Robinson (2011).

With the outbreak of COVID-19, the economics literature exploded. A large number of papers focus on the optimal control of the pandemic by policy-makers, given a trade-off between deaths from the disease and GDP losses from lockdowns. Alvarez, Argente, and Lippi (2021) is an example of this class of models in the case of homogeneous agents while Acemoglu, Chernozhukov, Werning, and Whinston (2021) and Favero, Ichino, and Rustichini (2020) focus on heterogeneous agents. Acemoglu et al. (2021) use a SAM-type matching function to model contacts and assume increasing returns to contacts.

Several other models solve for equilibrium in the decentralized economy, in which optimizing agents respond to economic incentives or infection disincentives, depending on their state and information set. The main states which are influenced by individual decision-making are the susceptible and infected. In the susceptible state the decision of the individual whether to interact in social space or not influences their probability of infection, whereas in the infected state the influence is on the probability of infecting others. These models often obtain also the planning solution. Several of the papers in this class of models use a variant of the matching function to derive contacts between agents, which are usually restricted to be either linear-homogeneous or quadratic. We discuss models in this class in the main body of the paper.[6]

Section 2 discusses pre-COVID research, with emphasis on HIV and influenza epidemics. Section 3 focuses on the large COVID-19 literature that exploded in 2020, distinguishing between models with homogenous agents and models with heterogeneous agents. We then discuss the role of externalities between the decentralized and centralized solutions in both pre-COVID and COVID research. Section 4 concludes.

2. PRE-COVID RESEARCH

We distinguish between two contrasting strands in the pre-COVID research. In the first strand authors addressed transmissions of HIV, or other sexually transmitted diseases (STDs), by making use of a SI epidemiological model; in other words, a model in which there are no deaths or recoveries. The disease imposes some cost on the infected, which gives the reason people want to avoid it. In STD models, the contact that leads to the transmission is planned and usually restricted to two people. This contrasts with the second class of models, the study of influenza epidemics, which usually makes use of the original SIR model, in which there are recoveries from the disease (Kermack & McKendrick, 1927). In this class of models, disease transmissions can also arise after unplanned chance meetings between individuals.

In both types of models, contact between infected and susceptible individuals has features of the SAM approach. The typical labor situation with SAM is one in which two agents come into contact with a view to forming a productive relationship. The match takes place if both parties agree to it. This parallels contacts in STD situations virtually exactly.[7] In contrast, infections in the case of influenza can arise in a variety of situations in which people share space, giving rise to a different set of solutions. For example, one might plan a restaurant visit and get infected by someone else who happens to be in the restaurant.

The implications for the contact technology are that in STD situations the contact function can be approximated by a linear one. For example, suppose that in an HIV world a person engages in sexual contact with n other persons and contact with each involves a disease transmission risk β. The infection probability for this person is $p = 1 - (1 - \beta)^n$. If β (and consequently p) is small, this approximately satisfies $p = \beta n$, a linear transmission rate.

But in influenza situations, as in COVID-19, there is a proportionally bigger effect on infections if a person increases or restricts social behavior. As an approximation, the contact technology is typically assumed to be quadratic. To see the origins of this, suppose a person goes into social space n times during the week. Each time she goes out she gets sufficiently close to m other people that could infect her with probability β each. If m is fixed and independent of n, we could reason as in the case of HIV and proxy the infection probability by $p = \beta nm$. But m must depend on the number of times other agents go to social space. Let $\overline{n}$ be the average number of times other agents go to social space, then m is proportional to $\overline{n}$, $m = x\overline{n}$; e.g., if everyone else doubles the number of times they go to social space, the person who goes out n times will be twice as likely to meet someone in that space. Therefore, the infection probability becomes $p = \beta x n \bar{n}$, and so in symmetric equilibrium, $p = \beta x n^2$.

In HIV situations, this "increasing returns" effect does not arise because if a person decides that they will have n partners, they can keep the n partners irrespective of how many partners others have and how many other "propositions" they get. But in influenza situations other people could affect the infections probability without the consent of the person in question.

We review each literature strand separately.

2.1 The SI Model and the Economics of HIV Epidemics

Early models (Geoffard & Philipson, 1996; Kremer, 1996) of STDs did not work with a general matching function in the spirit of SAM theory, but interpreted instead the standard incidence of the epidemiological model (Hetcote, 2000) through the lens of a random matching game. Of course, random matching is consistent with a simple matching function in which the aggregate probability of infection depends linearly on the stock of infected people, and this is the main transmission mechanism explored by these papers. The linear dependence of the transmission probability on the number of infected people is also a standard result in SAM models with linear technologies.

Consider the SI model studied by Toxvaerd (2019), who discusses the differences between the model without an economic dimension and the impact that economic incentives might have on it. In the former, a population $P = [0, 1]$ consists of a continuum of infinitely lived individuals who at each instant $t \geq 0$ can be in one of two states, namely, susceptible or infected. The set of infected individuals is denoted by $I(t)$ and has measure $I(t)$, whereas the set of susceptible individuals is denoted by $S(t)$ and has measure $S(t)$. In the absence of births and deaths the population size is normalized to unity, so these measures can be interpreted as fractions. $I(t)$ is referred to as the *disease prevalence*.[8]

At each instant, the population mixes homogeneously. This corresponds to *random matching*, where each individual has an equal chance of meeting any other individual, irrespective of the health status of the two matched individuals. A match between two infected individuals or two susceptible individuals does not create a new infection, but a match between an infected and a susceptible individual may do. In a continuous time model, the rate at which infection is transmitted in a match with a member of the $I(t)$ set is denoted by $\beta > 0$, so in a short interval of time dt the rate of getting the disease from an infected individual is βdt. This parameter captures the infectiousness of the disease. With random matching and large numbers, it follows that the probability that a susceptible person in this population is infected during the short interval dt is

$$\lambda(t) = \beta I(t) dt.$$

It follows that the average rate at which susceptible individuals become infected is

$$-\frac{dS(t)}{dt} = \beta I(t)S(t), \tag{1}$$

with $-\frac{dS(t)}{dt}$ giving the flow from the susceptible state to the infected. Thus, the rate of new infections, or *disease incidence*, is proportional to disease prevalence.

With this interpretation, the key transmission rate can be interpreted as the outcome of a SAM mechanism with linear technology (Diamond and Maskin, 1979). To turn the classical SI model into an economic model, Toxvaerd (2019) assumes that individuals earn a flow payoff $\pi_S > 0$ per instant while susceptible, a flow payoff $\pi_I < \pi_S$ per instant while infected, and that time is discounted at rate ρ. In Toxvaerd's notation $\pi = \pi_S - \pi_I > 0$ denotes the health premium. The health premium should be thought of broadly as the benefits of not being infected. To model the possibility of engaging in preventive behavior, assume that the individuals can affect the rate of infection by controlling the rate at which they expose themselves to a potential infection. In particular, at each instant $t \geq 0$, each susceptible individual $i \in S(t)$ noncooperatively chooses exposure level $\epsilon_i(t) \in [0, 1]$, at personal cost $(1 - \epsilon_i(t))c \geq 0$. Here, $\epsilon_i(t) = 0$ denotes complete shielding from social action whereas $\epsilon_i(t) = 1$ denotes no shielding at all, so c is the unit cost of shielding from social action. The introduction of shielding reduces the rate of infection to $\epsilon_i(t)\beta I(t)$. This formalization captures the notion that exposure is desirable, but shielding is pursued because it reduces the chance of an infection and the loss of the health premium. In a symmetric equilibrium, $\epsilon_i(t) = \epsilon(t)$ and the aggregate infection rate becomes (on the assumption that the infected population chooses full exposure),

$$-\frac{dS(t)}{dt} = \epsilon(t)\beta I(t)S(t).$$

Several new results follow from this general framework, since the choice of exposure depends both on the economic costs of contracting the disease and on the cost of shielding.

Returning now to HIV, in an early paper, Geoffard and Philipson (1996) show that in an optimizing model the aggregate transmission rate depends on incidence and derive some results about behavior from it. The model is SI; there is a population composed of agents who are either susceptible or infected, and engage in either protective or transmissive (exposed) activity. When susceptible, an agent can either become infected or remain susceptible; once infected, an agent remains infected. Agents continuously meet one another over time, and upon each meeting, they must decide whether to engage in transmissive or protective behavior. Susceptible agents who choose the former run the risk of contracting the disease, while susceptible agents who choose the latter run no such risk. Transmissive behavior is assumed to be desirable, and protective behavior costly. Since infection is an absorbing state, in the framework of Geoffard and Philipson (1996) no selfishly rational infected agent engages in protection and their problem is basically static. Even though individual behavior takes the form of a binary decision, because of large numbers at the aggregate level the hazard function from susceptible to infected depends continuously on individual choices.

Geoffard and Philipson (1996) argue that "in standard mathematical epidemiology, this hazard rate is an increasing function of prevalence. In other words, the larger the fraction of infected people in the population, the larger is the fraction of uninfected people who become infected in the next period. This is because the larger the disease is, the larger is the chance that an individual who is still susceptible will meet an infected individual. This is true across a wide variety of epidemiological models, since they all share the feature that the demand for exposure does not respond to prevalence." In contrast, introducing economic considerations into the model implies that the hazard rate into infection may be a decreasing function of the prevalence of the disease because the individuals who are still susceptible face a larger risk of infection and increase protective behavior. Geoffard and Philipson (1996) show that the aggregate dynamics of the model can be written as

$$-\frac{dS(t)}{dt} = \beta G(I(t))Q(t)I(t). \tag{2}$$

As before, $S(t)$ and $I(t)$ denote the susceptible and infected population, respectively, and β denotes the infectiousness of the disease. $Q(t)$ denotes the probability that a susceptible agent who engages in transmissive activity (in the case of HIV, engages in sexual activity) during period $[0, t]$ is still susceptible at t, and $G(I(t))$ is an endogenous probability that keeps track of the share of susceptible people who engage in transmissive activity in t, as a function of prevalence $I(t)$. The function $G(I(t))$ picks up the disincentives that susceptible agents have; unlike the earlier epidemiological models without disincentives from prevalence, in which $G(.) = I(t)$, Geoffard and Philipson (1996) show that $G(.)$ is a decreasing function of prevalence $I(t)$. Note that it is only the present level of prevalence that matters, not expected future prevalence levels. To understand why, note that $I(t)$ is increasing in t as there are no deaths in the model. Furthermore, protective behavior reduces the probability of infection to zero. As a result, the individual decision of whether to protect oneself or not becomes like an optimal stopping problem, and the person stops transmissive behavior when the instantaneous gain from this behavior is exactly equal to the instantaneous cost associated with the risk of becoming ill. After that point, the person will never again engage in risky behavior. The future development of $I(t)$ is therefore irrelevant for the optimal stopping decision.[9]

A matching interpretation for this mechanism is one in which an individual is engaged in search sequentially but stops the search and enters an absorbing state when the cost of continuing rises up to the benefits of stopping. In this case the cost of continuing is the risk of infection, which rises during search because of the monotonic increase in infections. As in the simple matching model, there is heterogeneity across individuals which is not modeled explicitly and not identified a priori, but reflects the frictions inherent in the matching function (Pissarides, 2000, pp. 3–4). The friction here is the information about the health status of other individuals in the market. When a susceptible individual meets another person, she has to decide if that person is a good match (susceptible) or a bad one (infected), and the more infected individuals there are, the more likely is the person to reject contacts. Of course, the absence of a priori information may lead to a failure early on, despite the low perceived probability $G(I_t)$.

Kremer (1996) – in a paper published in the same year as Geoffard and Philipson (1996) – also argues that most epidemiological models treat behavior as independent of prevalence. In contrast, he argues that his "analysis differs both from the traditional epidemiological analysis, which takes behavior as independent of prevalence, and from the few attempts to introduce behavioral considerations into epidemiology, which do not formally model how decisions about the rate of partner change depend on the composition of available partners." Kremer (1996) models the behavioral choice in the rate of partner change by writing

$$-\frac{dS(t)}{dt} = i(I(t))\beta I(t)S(t), \tag{3}$$

where the function $i()$ is the rate of partner change. The key contribution is thus similar to Geoffard and Philipson (1996), in the sense that the rate of partner change is the outcome of an optimizing decision that corresponds to partner selection. This rate depends on the stock of infection, referred to as incidence in this literature. The motivation given is different, and appealing in the context of HIV, in that partner change involves a potential cost, the possibility of getting infected from the new partner. In the context of social activity in epidemics that we discussed in our introduction, staying with the same partner in HIV epidemics has a similar impact on the spread of the disease as complete social distancing in epidemics such as COVID-19.[10]

Greenwood, Kircher, Santos, and Tertil (2019) introduce directed search ideas into the HIV model, by distinguishing three states in which agents may find themselves: healthy, infected, and infected with treatment.[11] The dynamics between the three states are simulated based on individual decisions that choose to participate in one of three alternative "markets," or "meeting places," respectively for single-partner long-term sex, casual sex with condoms, or casual unprotected sex. This segregation of market structures eliminates any complexities due to differences in the interests of partners: they have the same intentions when they enter the same market. Finding a partner generates utility from sexual behavior. Marriage has the additional benefit of continued interaction without the need to search again.

There are four types of status for each individual, and these are labeled abstinence, long-term sexual relationship, short-term unprotected sex, and short-term protected sex. In each of these states people can be susceptible to HIV, infected with no treatment and infected with treatment. The health status is only known to the individual and it cannot be observed by the sexual partners.[12] Infected and susceptible people choose rationally the submarket to enter by optimally choosing the odds of finding partners in each of the three sexual submarkets (π_l, π_p, π_u where subscript l, p, u refer respectively to the odds of finding