Biblioteca del Instituto de Estadística y Cartografía de Andalucía

Resúmenes de revistas
Septiembre - octubre 2020

Junta de Andalucía

Consejería de Transformación Económica, Industria, Conocimiento y Universidades

INSTITUTO DE ESTADÍSTICA Y CARTOGRAFÍA DE ANDALUCÍA

PRESENTACIÓN

El presente boletín de resúmenes tiene una periodicidad bimestral y con él la Biblioteca del Instituto de Estadística y Cartografía de Andalucía pretende dar a conocer a los usuarios de una forma detallada el contenido de las revistas especializadas que entran en su colección. Se trata de un complemento al boletín de novedades de publicaciones seriadas ya que en él se incluyen los resúmenes de cada uno de los artículos que aparecen publicados en los diferentes números de las revistas en el idioma original de las mismas.

Los resúmenes de este boletín corresponden a las revistas que han ingresado en la Biblioteca del Instituto de Estadística y Cartografía de Andalucía durante los meses de **septiembre y octubre de 2020** y que pueden consultarse gratuitamente en sus instalaciones en la siguiente dirección:

Instituto de Estadística y Cartografía de Andalucía
Pabellón de Nueva Zelanda
C/Leonardo Da Vinci, n. 21. Isla de La Cartuja
41071 - SEVILLA
E-mail: biblio.ieca@juntadeandalucia.es
Teléfono: 955 033 800
Fax: 955 033 816

Horario de atención al público:
Lunes y jueves: de 9:00h a 14:00h. y de 16:00 a 19:00 h.
Martes, miércoles y viernes: de 9:00h a 14:00h.
Horario de verano (del 15 de junio al 15 de septiembre), Semana Santa, Feria de Sevilla y
Navidad (del 24 de diciembre al 6 de enero): de lunes a viernes de 9:00h. a 14:00h.

The Evolution of Cadastral Systems in Austria and Galicia (Poland): Different Approaches to a Similar System from a Common Beginning

Józef Hernik , Barbara Prus , Robert Dixon-Gough , Jarosław Taszakowski , Gerhard Navratil & Reinfried Mansberger

Abstract

 The main aim of this paper is to document the gradual evolution of cadastral maps and associated land books in the area of today's Austria as well as the more dramatic development of the cadastral system in Galicia (Poland). The continuous development of the Austrian system is compared to the more gradual evolution of the Polish system, the development of which was more complex because of the historical and political developments in this country's turbulent past. However both systems have common roots with the Franciscan Cadastre. The investigation is based on development of the legal instruments, the institutional settings, the technical procedures, and the final products of the respected countries. The comparison of the development in the two countries and the results of the investigated quality assessment of cadastral maps clearly document that the smooth development of a cadastre requires a stable political situation. The Austrian cadastre had already almost finished the period of consolidation, harmonization, and completion in 1938. This allowed a new area of further development to evolve after 1945. In Poland, the unfinished tasks of homogenization and harmonization in 1939 and the significant changes in the political system in 1945 delayed consolidation, harmonization, and completion of the cadastral system and impeded the continuous development of the system for the next decades.

Scope of Information in the Legends of Topographical Maps in the Nineteenth Century – *Urmesstischblätter*

Dariusz Lorek & Beata Medyńska-Gulij

Abstract

The aim of this study is to determine the informational resource of maps (two folios) of the *Urmesstischblätter* with reference to the legend (three editions). Visual analysis was used as the basis for acquiring information that could be gleaned from the legends, contents of the folios, marginalia and later map. The information gathered was grouped applying the geoinformational method. The informational resource of the folios was determined by identifying separate objects present on the map with reference to the legend (number of separations in the legend), taking into consideration the geometrical type of objects. In addition, the informational resource of a folio was also defined by the number of features of these objects that can be obtained (from the legend, cartographic content, marginalia, *Messtischblätter*) with reference to cartographic measurement levels.

Exploring Evocative Places and their Characteristics

Alenka Poplin

Abstract

The main goal of this article is to study how people describe emotions related to places and to improve understanding of

evocative places, their locations and characteristics as described by the citizens. Evocative places are places that evoke an especially emotional response in terms of images, memories and/or emotions. We aim to contribute to a better understanding of emotions in relation to places. Our study concentrates on Ames, a small university town in the Midwest in North America. The collected 192 locations show two areas with high concentrations of evocative places, specifically the university campus and Ada Hayden Heritage Park. This study opened up challenging research questions related to fuzzy or non-existing boundaries of places, a variety of shapes of places, and questions related to the challenges of modelling emotions as objects or attributes in a geographic information system. We conclude the article with main findings and further research directions.

Cutting the Cord: A Corrective for World Navels in Cartography and Science

Rasmus Grønfeldt Winther

Abstract

A map is not its territory. Taking a map too seriously may lead to *pernicious reification*: map and world are conflated. As one family of cases of such reification, I focus on maps exuding the *omphalos syndrome*, whereby a centred location on the map is taken to be the *world navel* of, for instance, an empire. I build on themes from my book *When Maps Become the World*, in which I analogize scientific theories to maps, and develop the tools of *assumption archaeology* and *integration platforms*. Here I argue that excavating assumptions helps fill *cartographic silences*, showing the limitations of perspectives often at war. Furthermore, integrating perspectives permits resisting imperial *central* or *master images*. A worthwhile future project would be a repository of world-navel maps, critically annotated with cultural context and imperial information. Mutual understanding may result from such an integration platform, perhaps implemented online or in a museum.

Territory and Claims in the Antarctic Treaty Region: A Disquisition on Historical and Recent Developments

Robert Keith Headland

Abstract

Formal specifications of territorial claims over Antarctic regions south of 60° south latitude, the region under the ægis of the Antarctic Treaty made on 1st December 1959, have been significant from their origin, early in the twentieth century, until several of the conundrums they raised went into abeyance after the Treaty came into force in 1961. Nevertheless, the history of such claims has not been entirely simple and there have been significant subsequent events involving Ecuador and Norway. Changes after the Antarctic Treaty are noted and the history of earlier territorial specifications recapitulated. Documents are quoted in the original languages. Comparative data of territories are presented with summarized observations on accuracy of cartographic representations.

An Alternative to Desktop GIS? Evaluating the Cartographic and Analytical Capabilities of WebGIS Platforms for Teaching
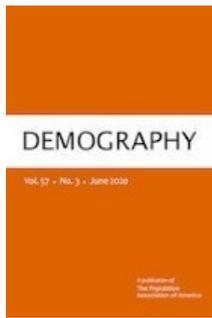
Victoria Fast & Fahim Hossain

Abstract

The increasing prevalence of (open) geospatial data has amplified the demand for online tools to input, manage, and analyse data and generate cartographic outputs. WebGIS platforms are responding to this demand by delivering cost-effective, relatively easy to use compared to desktop GIS software, and increasingly sophisticated tools. The continued advancement of webGIS platforms highlights the necessary, but tricky, integration of these platforms into university curriculum. In this paper, we develop an evaluation framework for users to assess, compare, and benchmark the cartographic capabilities of four webGIS platforms–ArcGIS Online, Carto, Mapbox, and GIS Cloud–against standard GIS functionality. The webGIS platforms studied offer advanced data management and some spatial analysis capabilities, but in contrast, they have limited cartographic capabilities for creating static maps. Recognizing that platforms change, and change often, this study serves to benchmark the capabilities of webGIS platforms, while reflecting on the observed challenges to teaching neocartography tools.

Abstract

The young adulthood years are demographically dense. Dr. Ronald Rindfuss made this claim when he was Population Association of America (PAA) president in 1991 (Rindfuss 1991), and this conclusion holds today. I offer both an update of his work by including Millennials and a new view on young adulthood by focusing on an increasingly common experience: cohabitation. I believe we need to move away from our marriage-centric lens of young adulthood and embrace the complexity that cohabitation offers. The cohabitation boom is continuing with no evidence of a slowdown. Young adults are experiencing complex relationship biographies, and social science research is struggling to keep pace. Increasingly, there is a decoupling of cohabitation and marriage, suggesting new ways of framing our understanding of relationships in young adulthood. As a field, we can do better to ensure that our theories, methods, and data collections better reflect the new relationship reality faced by young adults.

Abstract

A focus of research on short interpregnancy intervals (IPI) has been on young disadvantaged women whose births are likely to be unintended. Later initiation of family formation in the United States and other high-income countries points to the need to also consider a woman's attributes indicative of readiness for purposefully accelerated family formation achieved through short IPIs. We test for whether factors indicating "reproductive readiness"—including being married, being older, and having just had a first birth or a birth later than desired—predict a woman's non-use of contraception in the postpartum months. We also test for whether this contraceptive non-use results explicitly from wanting to become pregnant again. The data come from the 2012–2015 Pregnancy Risk Assessment Monitoring System, representing women who recently gave birth in any of 35 U.S. states and New York City ($N = 120,111$). We find that these reproductive-readiness factors are highly predictive of women's postpartum non-use of contraception because of a stated desire to become pregnant and are moderately predictive of contraceptive non-use without an explicit pregnancy intention. We conclude that planning for, or ambivalently risking, a short IPI is a frequent family-formation strategy for women whose family formation has been delayed. This is likely to become increasingly common as family formation in the United States is initiated later in the reproductive life course.

Abstract

In late 2015, the Brazilian Ministry of Health and the Pan American Health Organization classified the increase in congenital malformations associated with the Zika virus (ZIKV) as a public health emergency. The risk of ZIKV-related

congenital syndrome poses a threat to reproductive outcomes that could result in declining numbers of live births and potentially fertility. Using monthly microdata on live births from the Brazilian Information System on Live Births (SINASC), this study examines live births and fertility trends amid the ZIKV epidemic in Brazil. Findings suggest a decline in live births that is stratified across educational and geographic lines, beginning approximately nine months after the link between ZIKV and microcephaly was publicly announced. Although declines in total fertility rates were small, fertility trends estimated by age and maternal education suggest important differences in how Zika might have impacted Brazil's fertility structure. Further findings confirm the significant declines in live births in mid-2016 even when characteristics of the municipality are controlled for; these results highlight important nuances in the timing and magnitude of the decline. Combined, our findings illustrate the value of understanding how the risk of a health threat directed at fetuses has led to declines in live births and fertility.

The Examination of Diffusion Effects on Modern Contraceptive Use in Nigeria

David K. Guilkey, Veronica Escamilla, Ilene S. Speizer

Abstract

This study uses data gathered for an evaluation of a Bill & Melinda Gates Foundation–funded initiative designed to increase modern contraceptive use in select urban areas of Nigeria. When the initiative was conceived, the hope was that any positive momentum in the cities would diffuse to surrounding areas. Using a variety of statistical methods, we study three aspects of diffusion and their effects on modern contraceptive use: spread through mass communications, social learning, and social influence. Using a dynamic causal model, we find strong evidence of social multiplier effects through social learning. The results for social influence and spread through mass communications are promising, but we are unable to identify definitive causal impacts.

Abortion Reporting in the United States: An Assessment of Three National Fertility Surveys

Laura Lindberg, Kathryn Kost, Mia Zolna

Abstract

Despite its frequency, abortion remains a highly sensitive, stigmatized, and difficult-to-measure behavior. We present estimates of abortion underreporting for three of the most commonly used national fertility surveys in the United States: the National Survey of Family Growth, the National Longitudinal Survey of Youth 1997, and the National Longitudinal Study of Adolescent to Adult Health. Numbers of abortions reported in each survey were compared with external abortion counts obtained from a census of all U.S. abortion providers, with adjustments for comparable respondent ages and periods of each data source. We examined the influence of survey design factors, including survey mode, sampling frame, and length of recall, on abortion underreporting. We used Monte Carlo simulations to estimate potential measurement biases in relationships between abortion and other variables. Underreporting of abortion in the United States compromises the ability to study abortion—and, consequently, almost any pregnancy-related experience—using national fertility surveys.

Son Preference and Fertility Decisions: Evidence From Spatiotemporal Variation in Korea

Seik Kim, Sam-Ho Lee

Abstract

Using Korean data, this study investigates whether son-favoring ideas or the preference for sons affect fertility decisions. Son-favoring fertility behavior in Korea is of interest because the sex ratio at birth has recovered to a natural level after having been very skewed. To isolate the effects of the preference for sons from the effects of the surrounding environment, we compare the fertility behavior of individuals living in the same region but who were born in different regions or years. Exploiting the male-female gap in educational achievement at the parents' time and place of birth as exogenous variation in the 2000 Census Korea 2% sample, we find that the strength of son preference formed at an early age is associated with the strength of son-favoring fertility behavior as adults. Our results indicate that parents are more likely to have a third child if

they happen to have only daughters as their first two children. More importantly, this tendency is stronger if parents were born in a spatiotemporal region with more skewed gender gap in educational investment. These findings are robust against various alternative specifications, including endogenous migration issues.

When Did the Health Gradient Emerge? Social Class and Adult Mortality in Southern Sweden, 1813–2015

Tommy Bengtsson, Martin Dribe, Jonas Helgertz

Abstract

Across today's developed world, there is a clear mortality gradient by socioeconomic status for all ages. It is often taken for granted that this gradient was as strong—or even stronger—in the past when social transfers were rudimentary and health care systems were less developed. Some studies based on cross-sectional data have supported this view, but others based on longitudinal data found that this was not the case. If there was no gradient in the past, when did it emerge? To answer this question, we examine social class differences in adult mortality for men and women in southern Sweden over a 200-year period, using unique individual-level register data. We find a systematic class gradient in adult mortality emerging at ages 30–59 only after 1950 for women and after 1970 for men, and in subsequent periods also observable for ages 60–89. Given that the mortality gradient emerged when Sweden transitioned into a modern welfare state with substantial social transfers and a universal health care system, this finding points to lifestyle and psychosocial factors as likely determinants.

Understanding Trends in the Concentration of Infant Mortality Among Disadvantaged White and Black Mothers in the United States, 1983–2013: A Decomposition Analysis

Wen Fan, Liying Luo

Abstract

The United States compares unfavorably with other high-income countries in infant mortality, which recent literature has attributed to the poor birth outcomes among disadvantaged (i.e., unmarried and less-educated) mothers. Describing and decomposing the trend of the concentration of infant mortality among disadvantaged mothers thus provides important clues for improving birth outcomes. We develop the infant mortality disadvantage index (IMDI) to measure such concentration. Using the 1983–2013 Birth Cohort Linked Birth and Infant Death data, we show that although the IMDI—as a measure of mortality inequality—was persistently higher for Blacks than Whites, the trends were different between the two groups. The IMDI declined for Black women; for White women, however, it increased in the 1980s, then plateaued until the early 2000s, and declined thereafter. We then use Das Gupta's decomposition method to assess the contribution of five demographic/social factors (age, education, marriage, fertility, and infant mortality) to the IMDI trend. Nonmarital fertility among women with less than 12 years of education contributed most to Whites' changing IMDI; for Blacks, a shrinking proportion of the less-educated group and declines in infant mortality among disadvantaged mothers contributed to their declining IMDI. These findings explicate links between population-level compositional changes and infant mortality inequality.

The Long-Term Costs of Family Trajectories: Women's Later-Life Employment and Earnings Across Europe

Joanne S. Muller, Nicole Hiekel, Aart C. Liefbroer

Abstract

The "motherhood earnings penalty" is a well-established finding in many Western countries. However, a divide between mothers and nonmothers might oversimplify reality given that the family life course has diversified over the last decades. In addition, whether family choices have consequences for women's employment and earnings in later life is not well known, particularly in a comparative perspective. Using data on 50- to 59-year-old women from the Generations and Gender Programme, the British Household Panel Survey, and SHARELIFE for 22 European countries, we derive a typology of women's family trajectories and estimate its association with women's later-life employment and earnings. Whereas family

trajectory–related differences with regard to employment were relatively small, our findings reveal a clear, long-lasting family trajectory gradient in earnings. Childless women (with or without a partner) as well as single mothers had higher personal earnings than women whose family trajectories combined parenthood and partnership. Moreover, in societies in which reconciliation of work and family during midlife is less burdensome, labor market outcomes of women following different family trajectories converge. Our findings show that women's fertility and partnership behavior are inevitably interrelated and jointly influence employment and earning patterns until later in life. The results imply that promoting equal employment opportunities could have long-lasting effects on women's economic independence.

Cohort Trends in the Association Between Sibship Size and Educational Attainment in 26 Low-Fertility Countries

Seongsoo Choi, Riley Taiji, Christiaan Monden

Abstract

Children with many siblings have lower average educational attainment compared with children raised in smaller families, and this disadvantage by sibship size has been observed across many countries. We still know remarkably little, however, about how sibship size disadvantage has changed within countries and how such trends vary across countries. Using comparative data from 111 surveys from 26 low-fertility countries, we find an overall trend of growing sibship size disadvantage across cohorts in the majority of countries: between the 1931–1940 birth cohort and the 1971–1980 birth cohort, 16 of 26 countries showed a statistically significant increase in sibship size disadvantage in education, while only two countries showed a significant reduction in sibship size disadvantage. The disadvantage in years of education associated with having an additional sibling increased remarkably in post-socialist (0.3) and East Asian countries (0.34) and, to a lesser extent, Western European countries (0.2). In contrast, this disadvantage showed little change in Nordic countries (0.05) and even decreased in Anglo-Saxon countries (–0.11). We discuss explanations and implications of our comparative evidence in the context of the intergenerational transmission of education.

Gender Segregation, Occupational Sorting, and Growth of Wage Disparities Between Women

Felix Busch

Abstract

Average female wages in traditionally male occupations have steeply risen over the past couple of decades in Germany. This trend led to a new and substantial pay gap between women working in male-typed occupations and other women. I dissect the emergence of these wage disparities between women, using data from the German Socio-Economic Panel (1992–2015). Compositional change with respect to education is the main driver for growing inequality. Other factors are less influential but still relevant: marginal returns for several wage-related personal characteristics have grown faster in male-typed occupations. Net of individual-level heterogeneity, traditionally male occupations have also become more attractive because of rising returns to task-specific skills. Discrimination of women in typically male lines of work seems to have declined, too, which erased part of the wage penalty these women had previously experienced. In sum, I document changes in the occupational sorting behavior of women as well as shifts in occupation-level reward mechanisms that have had a profound impact on the state of inequality between working women.

Self-selection of Asylum Seekers: Evidence From Germany

Lucas Guichard

Abstract

I examine the pattern of selection on education of asylum seekers recently arrived in Germany from five key source countries: Afghanistan, Albania, Iraq, Serbia, and Syria. The analysis relies on original individual-level data collected in Germany combined with surveys conducted at origin. The results reveal a positive pattern of selection on education for asylum seekers who were able to flee Iraq and Syria, and the selection is neutral for individuals seeking asylum from Afghanistan and negative for asylum seekers from Albania and Serbia. I provide an interpretation of these patterns

based on differences in the expected length of stay at destination, the migration costs faced by asylum seekers to reach Germany, and the size of migration networks at destination.

Does Information Change Attitudes Toward Immigrants?
P. 1117–1143
Authors

Alexis Grigorieff, Christopher Roth, Diego Ubfal

Abstract

Strategies aimed at reducing negative attitudes toward immigrants are at the core of integration policies. A large literature shows that misperceptions about the size and characteristics of immigrants are common. A few studies implemented interventions to correct innumeracy regarding the size of the immigrant population, but they did not detect any effects on attitudes. We study whether providing information not only about the size but also about the characteristics of the immigrant population can have stronger effects. We conduct two online experiments with samples from the United States, providing one-half of the participants with five statistics about immigration. This information bundle improves people's attitudes toward current legal immigrants. Most effects are driven by Republicans and other groups with more negative initial attitudes toward immigrants. In our second experiment, we show that treatment effects persist one month later. Finally, we analyze a large cross-country survey experiment to provide external validity to the finding that information about the size of the foreign-born population is not enough to change policy views. We conclude that people with negative views on immigration before the intervention can become more supportive of immigration if their misperceptions about the characteristics of the foreign-born population are corrected.

From Malthusian Disequilibrium to the Post-Malthusian Era: The Evolution of the Preventive and Positive Checks in Germany, 1730–1870
P. 1145–1170

Ulrich Pfister, Georg Fertig

Abstract

This study draws on a new data set of vital rates and real wages to explore short-term and long-term behavior of the preventive and positive checks in a major economy of premodern mainland Europe. Four results stand out. First, the preventive check was fairly stable throughout the period 1730–1870; its magnitude of 0.2 to 0.35 was comparable with that of England, northern and central Italy, and Sweden. Second, the eighteenth century was characterized by Malthusian disequilibrium in that there was no long-term relationship between the crude death rate and the real wage, whereas the crude death rate's instantaneous response to income changes was a substantial –0.4. Third, the short-term positive check may have weakened over the eighteenth century and largely disappeared in the 1810s. The diversification of food risk resulting from the spread of potato cultivation, market integration, and the development of the nonagricultural sectors are potential explanations of the demise and disappearance of the positive check. Fourth, between the 1810s and the 1860s, vital rates and the real wage were stationary, which is consistent with a post-Malthusian regime in which technological progress depended on population size. The 1810s marked the time when Germany transited from a Malthusian regime in disequilibrium to the post-Malthusian era.

A Bayesian Reconstruction of a Historical Population in Finland, 1647–1850
P. 1171–1192

Miikka Voutilainen, Jouni Helske, Harri Högmander

Abstract

This article provides a novel method for estimating historical population development. We review the previous literature on historical population time-series estimates and propose a general outline to address the well-known methodological problems. We use a Bayesian hierarchical time-series model that allows us to integrate the parish-level data set and prior population information in a coherent manner. The procedure provides us with model-based posterior intervals for the final population estimates. We demonstrate its applicability by estimating the long-term development of Finland's population from 1647 onward and simultaneously place the country among the very few to have an annual population series of such length available.

Does Opportunity Skip Generations? Reassessing Evidence From Sibling and Cousin Correlations

Ian Lundberg

Abstract

Sibling (cousin) correlations are empirically straightforward: they capture the degree to which siblings' (cousins') socioeconomic outcomes are similar. At face value, these quantities seem to summarize something about how families constrain opportunity. Their meaning, however, is complicated. One empirical set of sibling and cousin correlations can be generated from a multitude of distinct theoretical processes. I illustrate this problem in the context of multigenerational mobility: the relationship between the incomes of an ancestor and a descendant separated by several generations in a family. When cousins' outcomes are similar (an empirical fact), prior authors have favored the particular theoretical interpretation that extended kin affect life chances through pathways not involving the parents of the focal individual. I show that this evidence is consistent with alternative theories of latent transmission (measurement error) or dynamic transmission (a parent-to-child transmission process that changes over generations). Theoretical assumptions are required to lend meaning to a point estimate. Further, I show that point estimates alone may be misleading because they can be highly uncertain. To facilitate uncertainty estimation for the key test statistic, I develop a Bayesian procedure to estimate sibling and cousin correlations. I conclude by outlining how future research might use sibling and cousin correlations as effective descriptive quantities while remaining cognizant that these quantities could arise from a variety of distinct theoretical processes.

The Emergence of Educational Hypogamy in India

Zhiyong Lin, Sonalde Desai, Feinian Chen

Abstract

With rising education among women across the world, educational hypergamy (women marrying men with higher education) has decreased over the last few decades in both developed and developing countries. Although a decrease in hypergamy is often accompanied by increasing homogamy (women marrying men with equal levels of education), our analyses for India based on a nationally representative survey of India (the India Human Development Survey), document a considerable rise in hypogamy (women marrying partners with lower education) during the past four decades. Log-linear analyses further reveal that declining hypergamy is largely generated by the rise in education levels, whereas hypogamous marriages continue to increase even after marginal distributions are taken into account. Further multivariate analyses show that highly educated women tend to marry men with lower education but from more privileged families. Moreover, consanguineous marriages, which exemplify strong cultural constraints on spousal selection in certain parts of India, are more likely to be hypogamous than marriages not related by blood. We argue that the rise in hypogamous marriage by education paradoxically reflects deep-rooted gender scripts in India given that other salient social boundaries are much more difficult to cross.

Educational Reproduction in Germany: A Prospective Study Based on Retrospective Data

Jan Skopek, Thomas Leopold

Abstract

This study examines educational reproduction of East and West German men and women born between 1930 and 1950. In a prospective design, we study the importance of mobility and fertility pathways of reproduction, considering not only the social reproduction of education as an attribute but also the demographic reproduction of individuals who carry this attribute. Using data from NEPS and SOEP, we introduce a method that estimates prospective models based on retrospective data commonly available in surveys. The analysis offers new estimates of the expected number of high- and low-educated children born to men and women of different levels of education. Findings show that the importance of the fertility pathway of educational reproduction was higher in West than in East Germany, higher for women than for men, and higher for earlier than for later cohorts. For West German women of earlier cohorts, the fertility pathway tempered educational reproduction among the high-educated and reinforced it among the low-educated. Population renewal models show that differential fertility slightly lowered educational attainment and slightly increased inequality in educational attainment in the offspring generation. Across cohorts, the fertility pathway declined in importance, a result of fertility convergence between education groups and educational expansion in postwar Germany. We conclude that prospective designs advance our understanding of educational reproduction. The method that we introduce substantially reduces the data requirements of prospective analysis, facilitating future prospective research on social stratification.

The Dynamics of U.S. Household Economic Circumstances Around a Birth

Alexandra B. Stanczyk

Abstract

With the arrival of an infant, many households face increased demands on resources, changes in the composition of income, and a potentially heightened risk of income inadequacy. Changing household economic circumstances around a birth have implications for child and family well-being, women's economic security, and public program design, yet have received little research attention in the United States. Using data from the Survey of Income and Program Participation, this study provides new descriptive evidence of month-to-month changes in household income adequacy and the composition of household income in the year before and after a birth. Results show evidence of significant declines in household income adequacy in the months around a birth, particularly for single mothers who live without other adults. Income from public benefit programs buffers but does not eliminate declines in income adequacy. Results have implications for policies targeted at this period, including public benefit and parental leave programs.

A Framework for Explaining Black-White Inequality in Homeownership Sustainability

Chunhui Ren

Abstract

To explain racially differential housing outcomes, previous studies have tended to concentrate on discriminatory processes within the mortgage market while ignoring homeowning families' broad socioeconomic challenges. This study proposes a conceptual framework for understanding Black-White inequality in homeownership sustainability, which emphasizes Black homeowners' socioeconomic challenges that are external to mortgage market evaluations, with a particular focus on the mediating role of liquid assets. Based on the Panel Study of Income Dynamics, the framework is put to an empirical test on the differential exit rates between Black and White homeowners in the United States during the recent housing crisis. The findings indicate that the racial gap in homeownership exit is eliminated after liquid wealth is controlled in the model alongside other covariates and that the inclusion of liquid wealth renders all mortgage-oriented variables nonsignificant with regard to their explanatory power for Black-White inequality in exit rates. Policy implications of the findings are also discussed.

Measuring Housing Stability With Consumer Reference Data

David C. Phillips

Abstract

Housing instability for low-income renters has drawn greater attention recently, but measurement has limited research on policies to stabilize housing. Address histories from consumer reference data can be used to increase the quantity and quality of research on low-income renters. Consumer data track housing moves throughout the entire United States for most of the adult population. In this article, I show that such data can measure housing stability for groups with very low income and extreme instability. For example, the data can track housing moves during natural disasters, at demolition of public housing, for households at high risk of homelessness, and during gentrification. Consumer data can track housing instability outcomes that are more common than shelter entry and less expensive to collect than surveys. Relative to existing administrative address histories, consumer data allow researchers to track housing moves to exact addresses and across jurisdictions.

Living Arrangements and Supplemental Income Programs for Older Adults in Mexico

Emma Aguila, Jung Ho Park, Alma Vega

Abstract

Living arrangements often reflect important quality-of-life indicators for elderly adults. In particular, increased income can prompt changes in household living arrangements for elderly adults. Using a differences-in-differences approach, we examine whether a supplemental income program in Mexico for adults aged 70 and older influenced household size and composition. We compare outcomes at baseline and at six-month follow-up for elderly adults in the treatment group with those in the control group that did not participate in the program. We find that household size increased by 3% in the treatment group relative to the control group. We also find a statistically significant increase in the number of girls aged 6–11 in the household, likely the granddaughters or great-granddaughters of program recipients. Increases in household size were greatest for adults aged 70–79, couples, households receiving two or more supplemental incomes, and households in the top income tercile. Household size did not increase for households of adults aged 80 and older, singles, households with only one supplemental income recipient, and households not in the top income tercile. These results suggest that when older adults have more income, they use part of this income to house their grandchildren.

The Household Structure Transition in China: 1982–2015

Ting Li, Wenting Fan, Jian Song

Abstract

Chinese society has experienced a dramatic change over the past several decades, which has had a profound impact on its household system. Utilizing the Chinese national census and 1% population survey data from 1982 to 2015, this study demonstrates the transition of the Chinese household structure through typology analyses. Five typical regional household structure types—large lineal, large nuclear, small nuclear, mixed lineal, and small and diverse—are identified. Our findings demonstrate that since the 1980s, the household system in almost all Chinese regions has evolved from a large unitary model to a small diversified one. However, this evolutionary path diverged after 2000 and formed two distinct household structure systems. There are also significant regional differences in the transition trajectory. Influenced by developmental, cultural, and demographic factors, the regions exhibit four distinct transition paths: lineal tradition, nuclear retardation, smooth transition, and fast transition. On the basis of these results, we discuss family modernization and other theories in explaining the transition of the Chinese household structure.

Counting on Potential Grandparents? Adult Children's Entry Into Parenthood Across European Countries

Roberta Rutigliano

Abstract

As populations age and longevity rises, the structure of the extended family is changing. Parents of young children are increasingly turning to the children's grandparents to provide childcare and help them reconcile work and family. This study is the first to investigate whether would-be grandparents' propensity to care for their grandchildren influences the adult children's transition to parenthood. Because grandparental childcare provision is not observable at the time of the transition to the first birth, I built a measure based on the characteristics of both actual grandparents and adult children to act as a proxy for the childcare that prospective grandparents are expected to provide in the future. Using data from the first two waves of the Survey of Health, Aging, and Retirement in Europe, I examine changes in the likelihood of having a first birth by different levels of expected future childcare provision. Given that the role grandparents play varies depending on the national context, I estimate distinct models for different groups of countries. Furthermore, I analyze different intensities of grandparental childcare: regular, occasional, and any other type of positive childcare. The comparison across 11 countries reveals that grandparental propensity to provide occasional childcare has a positive effect on the transition to parenthood in all country clusters and that grandparental propensity to provide regular childcare has a positive and significant association with having a first child in both pro-natalist (Belgium and France) and pro-traditional countries (Austria, Germany, Greece, Italy, Spain, and Switzerland).

Household Composition and Gender Differences in Parental Time Investments
P. 1415–1435

Andrew J. Bibler

Abstract

Recent research has documented the relatively poor performance of boys, especially those from single-mother households, on a number of outcomes. Differences in noncognitive skills are often cited as a main contributing factor. However, we still know little about the underlying mechanisms driving differences in noncognitive skills and other outcomes. This article provides empirical evidence that parental time investments, defined as the amount of time that parents spend participating in activities with their child, change differentially by child gender following a transition from a two-parent to single-mother household. Boys experience larger investment reductions following the change in household structure, which may help facilitate previously documented gender gaps in noncognitive skills for those in single-mother households. Boys lose an estimated additional 3.8 hours per week in fathers' time investments, nearly 30% of average weekly paternal investments across the sample. The difference is increasing with age, concentrated in leisure and entertainment activities, with little to no evidence that mothers increase investments in boys relative to girls after such transitions.

Evacuees and Migrants Exhibit Different Migration Systems After the Great East Japan Earthquake and Tsunami
P. 1437–1457

Mathew E. Hauer, Steven R. Holloway, Takashi Oda

Abstract

Research on the destinations of environmentally induced migrants has found simultaneous migration to both nearby and long-distance destinations, most likely caused by the comingling of evacuee and permanent migrant data. Using a unique data set of separate evacuee and migration destinations, we compare and contrast the pre-, peri-, and post-disaster migration systems of permanent migrants and temporary evacuees of the Great East Japan Earthquake and Tsunami. We construct and compare prefecture-to-prefecture migration matrices for Japanese prefectures to investigate the similarity of migration systems. We find evidence supporting the presence of two separate migration systems—one for evacuees, who seem to emphasize short distance migration, and one for more permanent migrants, who emphasize migration to destinations with preexisting ties. Additionally, our results show that permanent migration in the peri- and post-periods is largely identical to the preexisting migration system. Our results demonstrate stability in migration systems concerning migration after a major environmental event.

The Intergenerational Impact of Terror: Did the 9/11 Tragedy Impact the Initial Human Capital of the Next Generation?

Ryan Brown

Abstract

Given the unexpected nature of the terrorist attacks of September 11, 2001, a specific cohort of children were exogenously exposed to increased maternal psychological stress *in utero*. Rich administrative data and the precise timing of the event allow this study to uniquely provide insights into the health effects of exposure to maternal psychological stress across gestation. Results suggest that children exposed *in utero* were born significantly smaller and earlier than previous cohorts. The timing of the effect provides evidence that intrauterine growth is specifically restricted by first trimester exposure to stress; reductions in gestational age and increases in the likelihood of being born at low (<2,500 grams) or very low (<1,500 grams) birth weight are induced by increased maternal psychological stress mid-pregnancy. This study also documents a positively selected post-attack fertility response, which would bias an evaluation that includes cohorts conceived after September 11, 2001, in the control group.

Interdependencies in Mothers' and Daughters' Work-Family Life Course Trajectories: Similar but Different?

Sergi Vidal, Philipp M. Lersch, Karsten Hank

Abstract

Women's life courses underwent substantial changes in the family and work domains in the second half of the twentieth century. The associated fundamental changes in opportunity structures and values challenged the importance of families of origin for individual life courses, but two research strands suggest enduring within-family reproduction of women's family behavior and work outcomes. We revisit this issue by studying two complementary types of intergenerational associations in women's combined work-family trajectories. On the one hand, we examine similarities across mothers' and daughters' work-family trajectories to address the direct within-family reproduction of female life courses (intergenerational persistence). On the other hand, we examine systematic associations between work-family trajectories that are typical in each generation to address intergenerational interdependencies beyond direct reproduction that account for individual and societal constrains and opportunities that each generation faced (intergenerational correspondence). We use a within-dyad approach to sequence analysis and examine combined work-family trajectories between ages 18 and 35 of two generations of women, born in 1930–1949 and in 1958–1981, within the same family drawn from the German Socio-Economic Panel. Overall, we find evidence of small but nontrivial persistence in work-family trajectories across generations that is partly attributed to within-family mechanisms of reproduction. In addition, we find correspondence across typical trajectory patterns of each generation, without daughters necessarily resembling their mothers' trajectories. The strength of the intergenerational associations varies by social background. Our research improves and broadens our understanding of the reproduction of female life courses across generations.

The Effects of Education on Mortality: Evidence From Linked U.S. Census and Administrative Mortality Data

Andrew Halpern-Manners, Jonas Helgertz, Evan Roberts

Abstract

Does education change people's lives in a way that delays mortality? Or is education primarily a proxy for unobserved endowments that promote longevity? Most scholars conclude that the former is true, but recent evidence based on Danish twin data calls this conclusion into question. Unfortunately, these potentially field-changing findings—that obtaining additional schooling has no independent effect on survival net of other hard-to-observe characteristics—have not yet been subject to replication outside Scandinavia. In this article, we produce the first U.S.-based estimates of the effects of education on mortality using a representative panel of male twin pairs drawn from linked complete-count census and death records. For comparison purposes, and to shed additional light on the roles that neighborhood, family, and genetic factors play in confounding associations between education and mortality, we also produce parallel estimates of the education-mortality relationship using data on (1) unrelated males who lived in different neighborhoods during childhood, (2) unrelated males

who shared the same neighborhood growing up, and (3) non-twin siblings who shared the same family environment but whose genetic endowments vary to a greater degree. We find robust associations between education and mortality across all four samples, although estimates are modestly attenuated among twins and non-twin siblings. These findings—coupled with several robustness checks and sensitivity analyses—support a causal interpretation of the association between education and mortality for cohorts of boys born in the United States in the first part of the twentieth century.

Revisiting the Fertility Transition in England and Wales: The Role of Social Class and Migration

Hannaliis Jaadla, Alice Reid, Joseph Day

Abstract

We use individual-level census data for England and Wales for the period 1851–1911 to investigate the interplay between social class and geographical context determining patterns of childbearing during the fertility transition. We also consider the effect of spatial mobility or lifetime migration on individual fertility behavior in the early phases of demographic modernization. Prior research on the fertility transition in England and Wales has demonstrated substantial variation in fertility levels and declines by different social groups; however, these findings were generally reported at a broad geographical level, disguising local variation and complicated by residential segregation along social class and occupational lines. Our findings confirm a clear pattern of widening social class differences in recent net fertility, providing strong support for the argument that belonging to a certain social group was an important determinant of early adoption of new reproductive behavior in marriage in England and Wales. However, a relatively constant effect of lower net fertility among long-distance migrants both before the transition and in the early phases of declining fertility indicates that life course migration patterns were most likely factor in explaining the differences in fertility operating through postponement of marriage and childbearing.

Twins Support the Absence of Parity-Dependent Fertility Control in Pretransition Populations

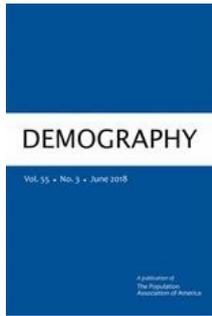Gregory Clark, Neil Cummins, Matthew Curtis

Abstract

A conclusion of the European Fertility Project in 1986 was that pretransition populations mostly displayed *natural fertility*, where parity-dependent birth control was absent. This conclusion has recently been challenged for England by new empirical results and has also been widely rejected by theorists of long-run economic growth, where pre-industrial fertility control is integral to most models. In this study, we use the accident of twin births to show that for three Western European–derived pre-industrial populations—namely, England (1730–1879), France (1670–1788), and Québec (1621–1835)—we find no evidence for parity-dependent control of marital fertility. If a twin was born in any of these populations, family size increased by 1 compared with families with a singleton birth at the same parity and mother age, with no reduction of subsequent fertility. Numbers of children surviving to age 14 also increased. Twin births also show no differential effect on fertility when they occurred at high parities; this finding is in contrast to populations where fertility is known to have been controlled by at least some families, such as in England, 1900–1949, where a twin birth increased average births per family by significantly less than 1.

The Effects of Enhanced Enforcement at Mexico's Southern Border: Evidence From Central American Deportees

Fernanda Martínez Flores

Abstract

Immigration enforcement cooperation between final-destination and transit countries has increased in the last decades. I examine whether the Southern Border Plan, an immigration enforcement program implemented by the Mexican government in 2014, has curbed intentions of unauthorized migrants from El Salvador, Guatemala, and Honduras to migrate to the United States. I use the announcement of the Southern Border Plan to implement a difference-in-differences approach and compare the evolution of short-run intentions to engage in additional unauthorized crossings of Central American (treatment group) relative to Mexican deportees (comparison group). The findings suggest that increased enforcement in Mexico decreases the likelihood of attempting repeated unauthorized crossings.

New Partner, New Order? Multipartnered Fertility and Birth Order Effects on Educational Achievement

Mats Lillehagen, Martin Arstad Isungset

Abstract

A substantial amount of research shows that younger siblings perform worse than their older sisters and brothers in several socioeconomic outcomes, including educational achievement. Most of these studies examined stable families and excluded half-siblings. However, the increasing prevalence of multipartnered fertility implies that many children grow up in nonnuclear families. We examine whether there is evidence for birth order effects in this context, which offers an opportunity to test and potentially expand the explanatory scope of the two main theories on birth order effects. We use comprehensive Norwegian registry data to study siblings in the 1985–1998 cohorts born to mothers or fathers who parented children with at least two partners. We provide evidence for negative effects of birth order on lower secondary school grades in both cases. Children born to fathers displaying multipartnered fertility tend to have lower grades than older full siblings but perform more similarly or better compared with older half-siblings. For siblings born to mothers with the multipartnered fertility pattern, later-born siblings do worse in school compared with all older siblings. This indicates that negative birth order effects tend to operate either within or across sets of full siblings, depending on the sex of the parent displaying multipartnered fertility. We argue that these findings can be explained by a combination of resource dilution/confluence theory and sex differences in residential arrangements following union dissolutions. We also suggest an alternative interpretation: maternal resources could be more important for generating negative birth order effects.

Brazil's Missing Infants: Zika Risk Changes Reproductive Behavior

Marcos A. Rangel, Jenna Nobles, Amar Hamoudi

Abstract

Zika virus epidemics have potential large-scale population effects. Controlled studies of mice and nonhuman primates

indicate that Zika affects fecundity, raising concerns about miscarriage in human populations. In regions of Brazil, Zika risk peaked months before residents learned about the epidemic and its relation to congenital anomalies. This spatiotemporal variation supports analysis of both biological effects of Zika infection on fertility and the effects of learning about Zika risk on reproductive behavior. Causal inference techniques used with vital statistics indicate that the epidemic caused reductions in birth cohort size of approximately one-quarter 18 months after Zika infection risk peaked but 10 months after public health messages advocated childbearing delay. The evidence is consistent with small but not statistically detectable biological reductions in fecundity, as well as large strategic changes in reproductive behavior to temporally align childbearing with reduced risk to infant health. The behavioral effects are larger for more-educated and older women, which may reflect facilitated access to information and to family planning services within high-risk, mosquito-infested urban locations as well as perceptions about the opportunity costs of risks to pregnancy and infant survival.

How Does Deprivation Affect Early-Age Mortality? Patterns of Socioeconomic Determinants of Neonatal and Postneonatal Mortality in Bolivia

Filippo Temporin

Abstract

Three mechanisms related to household living standards might affect early-age mortality: the absolute level of deprivation, its level relative to the average of the community, and the inequality in the distribution of deprivation within communities. A large body of literature has explored the effect of the absolute level of deprivation, but little research has examined the association between relative deprivation and early-age mortality, and findings related to deprivation inequality are inconsistent. Using 2008 Bolivian Demographic and Health Survey data, this study explores patterns of association between the three factors and mortality occurring in the neonatal and postneonatal periods. Because household-level deprivation might capture some unmeasured characteristics at the community level, such as area-specific investments, this study decomposes household-level deprivation into its between- and within-community components. The results show that after possible confounders are controlled for, community-level absolute deprivation is a significant predictor of neonatal and postneonatal mortality. Relative deprivation and deprivation inequality are not associated with early-age mortality. These findings are specific to a context of widespread deprivation and low inequality within communities; the role of the distribution of deprivation might be more important in countries in which basic needs are met within a bigger proportion of the population. This study helps identify crucial sectors of development related to living standards and deprivation inequality in order to tackle neonatal and postneonatal mortality.

Biases in Survey Estimates of Neonatal Mortality: Results From a Validation Study in Urban Areas of Guinea-Bissau

Stéphane Helleringer, Li Liu, Ane Barent Fisker

Abstract

Neonatal deaths (occurring within 28 days of birth) account for close to one-half of all deaths among children under age 5 worldwide. In most low- and middle-income countries, data on neonatal deaths come primarily from household surveys. We conducted a validation study of survey data on neonatal mortality in Guinea-Bissau (West Africa). We used records from an urban health and demographic surveillance system (HDSS) that monitors child survival prospectively as our reference data set. We selected a stratified sample of 599 women aged 15–49 among residents of the HDSS and collected the birth histories of 422 participants. We cross-tabulated survey and HDSS data. We used a mathematical model to investigate biases in survey estimates of neonatal mortality. Reporting errors in survey data might lead to estimates of the neonatal mortality rate that are too high, which may limit our ability to track progress toward global health objectives.

The Social Significance of Interracial Cohabitation: Inferences Based on Fertility Behavior

Kate H. Choi, Rachel E. Goldberg

Abstract

Interracial couples cohabit at higher rates than same-race couples, which is attributed to lower barriers to interracial cohabitation relative to intermarriage. This begs the question of whether the significance of cohabitation differs between interracial and same-race couples. Using data from the 2006–2017 National Survey of Family Growth, we assessed the meaning of interracial cohabitation by comparing the pregnancy risk, pregnancy intentions, and union transitions following a pregnancy among women in interracial and same-race cohabitations. The pregnancy and union transition behaviors of women in White-Black cohabitations resembled those of Black women in same-race cohabitations, suggesting that White-Black cohabitation serves as a substitute to marriage and reflecting barriers to the formation of White-Black intermarriages. The behaviors of women in White-Hispanic cohabitations fell between those of their same-race counterparts or resembled those of White women in same-race cohabitations. These findings suggest that White-Hispanic cohabitations take on a meaning between trial marriage and substitute to marriage and support views that Hispanics with White partners are a more assimilated group than Hispanics in same-race unions. Results for pregnancy intentions deviated from these patterns. Women in White-Black cohabitations were less likely than Black women in same-race cohabitations to have an unintended pregnancy, suggesting that White-Black cohabitations are considered marriage-like unions involving children. Women in White-Hispanic cohabitations were more likely than White and Hispanic women in same-race cohabitations to have an unintended pregnancy, reflecting possible concerns about social discrimination. These findings indicate heterogeneity in the significance of interracial cohabitation and continuing obstacles to interracial unions.

Marriage and Union Formation in the United States: Recent Trends Across Racial Groups and Economic Backgrounds

Deirdre Bloome, Shannon Ang

Abstract

Family formation in the United States has changed dramatically: marriage has become less common, nonmarital cohabitation has become more common, and racial and economic inequalities in these experiences have increased. We provide insights into recent U.S. trends by presenting cohort estimates for people born between 1970 and 1997, who began forming unions between 1985 and 2015. Using Panel Study of Income Dynamics data, we find that typical ages at marriage and union formation increased faster across these recent cohorts than across cohorts born between 1940 and 1969. As fewer people married at young ages, more cohabited, but the substitution was incomplete. We project steep declines in the probability of ever marrying, declines that are larger among Black people than White people. We provide novel information on the intergenerational nature of family inequalities by measuring parental income, wealth, education, and occupational prestige. Marriage declines are particularly steep among people from low-income backgrounds. Black people are overrepresented in this low-income group because of discrimination and opportunity denial. However, marriage declines are larger among Black people than White people across parental incomes. Further, most racial differences in marriage occur among people from similar socioeconomic backgrounds. Family inequalities increasingly reflect both economic inequalities and broader racial inequalities generated by racist structures; in turn, family inequalities may prolong these other inequalities across generations.

The Direct Effects of Legal Same-Sex Marriage in the United States: Evidence From Massachusetts

Christopher S. Carpenter

Abstract

I provide evidence on the direct effects of legal same-sex marriage in the United States by studying Massachusetts, the first state to legalize it in 2004 by court order. Using confidential Massachusetts data from 2001–2013, I show that the ruling significantly increased marriage among lesbians, bisexual women, and gay men compared with the associated change for heterosexuals. I find no significant effects on coupling. Marriage take-up effects are larger for lesbians than for bisexual women or gay men and are larger for households with children than for households without children. Consistent with prior work in the United States and Europe, I find no reductions in heterosexual marriage.

The Wealth of Parents: Trends Over Time in Assortative Mating Based on Parental Wealth

Sander Wagner, Diederik Boertien, Mette Gørtz

Abstract

This article describes trends in parental wealth homogamy among union cohorts formed between 1987 and 2013 in Denmark. Using high-quality register data on the wealth of parents during the year of partnering, we show that the correlation between partners' levels of parental wealth is considerably lower compared with estimates from research on other countries. Nonetheless, parental wealth homogamy is high at the very top of the parental wealth distribution, and individuals from wealthy families are relatively unlikely to partner with individuals from families with low wealth. Parental wealth correlations among partners are higher when only parental assets rather than net wealth are examined, implying that the former might be a better measure for studying many social stratification processes. Most specifications indicate that homogamy increased in the 2000s relative to the 1990s, but trends can vary depending on methodological choices. The increasing levels of parental wealth homogamy raise concerns that over time, partnering behavior has become more consequential for wealth inequality between couples.

The Direct Effect of Taxes and Transfers on Changes in the U.S. Income Distribution, 1967–2015

Christopher Wimer, Zachary Parolin, Christopher Jencks

Abstract

Scholars have increasingly drawn attention to rising levels of income inequality in the United States. However, prior studies have provided an incomplete account of how changes to specific transfer programs have contributed to changes in income growth across the distribution. Our study decomposes the direct effects of tax and transfer programs on changes in the household income distribution from 1967 to 2015. We show that despite a rising Gini coefficient, lower-tail inequality (the ratio of the 50th to 10th percentile) *declined* in the United States during this period due to the rise of in-kind and tax-based transfers. Food assistance and refundable tax credits account for nearly all the income growth between 1967 and 2015 at the 5th percentile and roughly one-half the growth at the 10th percentile. Moreover, income gains near the bottom of the distribution are concentrated among households with children. Changes in the income distribution were far less progressive among households without children.

Early Determinants of Work Disability in an International Perspective

Axel Börsch-Supan, Tabea Bucher-Koenen, Felizia Hanemann

Abstract

This study explores the interrelated roles of health and welfare state policies in the decision to take up disability insurance (DI) benefits due to work disability (WD), defined as the (partial) inability to engage in gainful employment as a result of physical or mental illness. We exploit the large international variation of health, self-reported WD, and the uptake of DI benefits in the United States and Europe using a harmonized data set with life history information assembled from SHARE, ELSA, and HRS. We find that the mismatch between WD and DI benefit receipt varies greatly across countries. Objective health explains a substantial share of the within-country variation in DI, but this is not the case for the variation across countries. Rather, most of the variation between countries and the mismatches are explained by differences in DI policies.

What Factors Explain the Decline in Widowed Women's Poverty?

Alicia H. Munnell, Geoffrey Sanzenbacher, Alice Zulkarnain

Abstract

Historically, women in widowhood in the United States have been vulnerable, with high rates of poverty. However, over the past several decades, their poverty rate has fallen considerably. In this article, we look at why this decline occurred and whether it will continue. Using data from the Health and Retirement Study linked to Social Security administrative earnings and benefit records, we address these questions by exploring three factors that could have contributed to this decline: (1)

women's rising levels of education; (2) their increased attachment to the labor force; and (3) increasing marital selection, reflecting that whereas marriage used to be equally distributed, it is becoming less common among those with lower socioeconomic status. The project decomposes the share of the decline in poverty into contributions by each of these factors and also projects the role of these factors in the future. The results indicate that increases in education and work experience have driven most of the decline in widows' poverty to date, but that marital selection will likely play a large role in a continuing decline in the future. Still, even after these effects play out, poverty among widows will remain well above that of married women.

Tapped Out? Racial Disparities in Extrahousehold Kin Resources and the Loss of Homeownership

Gregory Sharp, Ellen Whitehead, Matthew Hall

Abstract

Research shows that extrahousehold kin economic resources contribute to the racial gap in transitions into homeownership, but the extent to which these resources matter for racial disparities in exits from homeownership is less understood. Using longitudinal data from the Panel Study of Income Dynamics, 1984–2017, we examine the role of extrahousehold kin wealth and poverty in shaping racial inequalities in the risk of exiting homeownership. Our nonlinear decomposition results indicate that racial differences in family network resources explain a nontrivial portion of the racial gap in homeownership exit, but there is little evidence that the effects of kin resources on exit are moderated by race. Among both Black and White owners, having wealthier noncoresident kin does not lessen the negative impacts of adverse economic or health shocks on the probability of losing homeownership. Our findings have implications for policies and programs designed to buttress the ability of minority households, especially those in financial distress, to sustain the wealth-building state of homeownership.

Metropolitan Reclassification and the Urbanization of Rural America

Kenneth M. Johnson, Daniel T. Lichter

Abstract

Research shows that extrahousehold kin economic resources contribute to the racial gap in transitions into homeownership, but the extent to which these resources matter for racial disparities in exits from homeownership is less understood. Using longitudinal data from the Panel Study of Income Dynamics, 1984–2017, we examine the role of extrahousehold kin wealth and poverty in shaping racial inequalities in the risk of exiting homeownership. Our nonlinear decomposition results indicate that racial differences in family network resources explain a nontrivial portion of the racial gap in homeownership exit, but there is little evidence that the effects of kin resources on exit are moderated by race. Among both Black and White owners, having wealthier noncoresident kin does not lessen the negative impacts of adverse economic or health shocks on the probability of losing homeownership. Our findings have implications for policies and programs designed to buttress the ability of minority households, especially those in financial distress, to sustain the wealth-building state of homeownership.

Income Segregation: Up or Down, and for Whom?

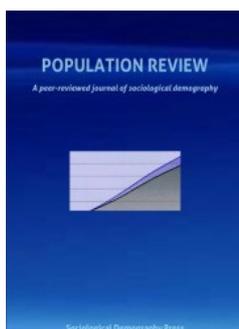John R. Logan, Andrew Foster, Wenquan Zhang

Abstract

Reports of rising income segregation in the United States have been brought into question by the observation that post-2000 estimates are upwardly biased because of a reduction in the sample sizes on which they are based. Recent studies have offered estimates of this sample-count bias using public data. We show here that there are two substantial sources of systematic bias in estimating segregation levels: bias associated with sample size and bias associated with using weighted sample data. We rely on new correction methods using the original census sample data for individual households to provide more accurate estimates. Family income segregation rose markedly in the 1980s but only selectively after 1990. For some categories of families, segregation declined after 1990. There has been an upward trend for families with children but not specifically for families with children in the upper or lower 10% of the income distribution. Separate analyses by race/ethnicity show that income segregation was not generally higher among Blacks and Hispanics than among White

families, and evidence of income segregation trends for these separate groups is mixed. Income segregation increased for all three racial groups for families with children, particularly for Hispanics (but not Whites or Blacks) in the upper 10% of the income distribution. Trends vary for specific combinations of race/ethnicity, presence of children, and location in the income distribution, offering new challenges for understanding the underlying processes of change.

Who's Skeptical of Vaccines? Prevalence and Determinants of Anti-Vaccination Attitudes in Italy

Ceylan Engin, Cristiano Vezzoni

Abstract

This paper investigates negative attitudes toward vaccines in Italy, where anti-vaccination movements have gained significant momentum in recent years. Considering the substantial health risk to herd immunity the issue poses, particularly after the sudden outbreak of the Covid-19 pandemic, it has become increasingly urgent to study the prevalence and diffusion of anti-vaccination beliefs. Using data from the 2016 European Social Survey's (ESS) country specific questions for Italy, the prevalence of anti-vaccination attitudes is examined along with how they are influenced by demographic, attitudinal and value-oriented determinants. The results show that 15 percent of the Italian public strictly hold negative views toward vaccination, and the prevalence of anti-vaccination attitudes is most commonly found among those who are less educated and aged between 25 and 34. While religiosity and political conservatism do not have an effect on anti-vaccination beliefs, our results indicate a strong positive link between anti-vaccination attitudes and distrust in the country's health-care systems and political institutions.

Early Childhood Parenting Practices in Indonesia

Mardiana Dwi Puspitasari, Aditya Rahmadhony, Sabarinah Prasetyo, Wisnu Fadila

Abstract

Parenting practices play an important role in early childhood development. Some literature and ethnographic studies suggest that collectivism, as part of local culture, is crucial in determining parenting practices in Indonesia. However, an inter-cultural study shows that parents' socioeconomic status has a stronger effect on the matter. The present study aims to examine the determinants of parenting practices in Indonesia, using data from the 2018 Population, Family Planning and Family Development Program Performance and Accountability Survey, a nation-wide survey conducted by the Indonesia National Population and Family Planning Board, with samples of 19,568 mothers of reproductive age who have children under 6 years of age. Parenting practices are measured using 17 items, covering both physical and psychosocial stimulation. Logistic regression was used to determine the predictors of early childhood parenting practices. The findings highlight that early childhood parenting practices are strongly associated with both cultural values and household wealth status. This study underscores the emergence of emotional/psychological interdependence values in Indonesia, characterised by high family integrity, and encouragement to a child's autonomy and achievement. Our model proposes a family's economic development and a maximum of two parity, which is important for Indonesian early childhood parenting practices.

Gender Role Attitudes and Fertility Revisited: Evidence from the United States

Qianqian Shang, Yongkun Yin

Abstract

Whether gender egalitarianism is associated with higher or lower fertility intentions is debatable. Some studies show

that gender egalitarianism is associated with higher fertility intentions; others document the opposite. Moreover, the interrelationship may vary by gender and across countries. Based on longitudinal data from the National Longitudinal Survey of Youth 1979 in the United States, we examine the effects of gender role attitudes with multiple measures of fertility (ideals, intentions and outcomes) and investigate how gender role attitudes are reshaped by the number of children. Our results show that individuals holding more egalitarian attitudes tend to have a smaller ideal family size, desire fewer children and have fewer children. In addition, for both men and women, the arrival of children can shape gender role attitudes towards more traditional ones. Different dimensions of gender attitudes may, however, affect and be affected by fertility to different extents, and the interrelationship can vary across gender. Our study adds more evidence to the debate over the effect of gender role attitudes on fertility, helps to understand distinct findings in the literature, sheds light on the development of gender role attitudes of men and women over time, and highlights the importance of using longitudinal data to examine the effects of gender attitudes on fertility behaviors.

Determinants and Spatial Patterns of Counterurbanization in Times of Crisis: Evidence from Greece

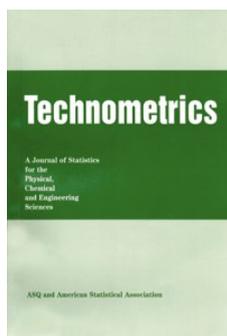Evgenia Anastasiou, Marie-Noëlle Duquenne

Abstract

The main purpose of this paper stems from the need for a systematic study of the multiple components that determine the factors that attract residents of urban centers to rural areas in Greece. Based on Multicriteria Analysis (Explanatory Factor Analysis and Hierarchical Analysis) of the last censuses data (2001–2011), the main types (spatial patterns) of Greek municipalities that have potential prospects for settlement are assessed. At the same time, age profiles of people who tend to enter specific types of municipalities are explored, highlighting both the attractiveness factors of the rural areas and what the internal migrant is looking for at the settlement destination. Distinct spatial patterns of counterurbanization can be identified in Greece based on age, physical amenities, employment structures, offer of services, degree of isolation, and the cost of living in the recipient location. Finally, the results identify that spatial, demographic, social and economic inequalities are decisive in the interpretation of internal migration flows.

A Latent Variable Approach to Gaussian Process Modeling with Qualitative and Quantitative Factors

P. 291-302

Yichi Zhang, Siyu Tao, Wei Chen & Daniel W. Apley

Abstract

Computer simulations often involve both qualitative and numerical inputs. Existing Gaussian process (GP) methods for handling this mainly assume a different response surface for each combination of levels of the qualitative factors and relate them via a multiresponse cross-covariance matrix. We introduce a substantially different approach that maps each qualitative factor to underlying numerical latent variables (LVs), with the mapped values estimated similarly to the other correlation parameters, and then uses any standard GP covariance function for numerical variables. This provides a parsimonious GP parameterization that treats qualitative factors the same as numerical variables and views them as affecting the response via similar physical mechanisms. This has strong physical justification, as the effects of a qualitative factor in any physics-based simulation model must *always* be due to some underlying numerical variables. Even when the underlying variables are many, sufficient dimension reduction arguments imply that their effects can be represented by a low-dimensional LV. This conjecture is supported by the superior predictive performance observed across a variety of examples. Moreover, the mapped LVs provide substantial insight into the nature and effects of the qualitative factors.

The Statistical Filter Approach to Constrained Optimization

P. 303-312

Tony Pourmohamad & Herbert K. H. Lee

Abstract

Expensive black box systems arise in many engineering applications but can be difficult to optimize because their output functions may be complex, multi-modal, and difficult to understand. The task becomes even more challenging when the optimization is subject to multiple constraints and no derivative information is available. In this article, we combine response surface modeling and filter methods in order to solve problems of this nature. In employing a filter algorithm for solving constrained optimization problems, we establish a novel probabilistic metric for guiding the filter. Overall, this hybridization of statistical modeling and nonlinear programming efficiently utilizes both global and local search in order to quickly converge to a global solution to the constrained optimization problem. To demonstrate the effectiveness of the proposed methods, we perform numerical tests on a synthetic test problem, a problem from the literature, and a real-world hydrology computer experiment optimization problem.

Bayesian Nonparametric Joint Mixture Model for Clustering Spatially Correlated Time Series

P. 313-329

Youngmin Lee & Heeyoung Kim

Abstract

We develop a Bayesian nonparametric joint mixture model for clustering spatially correlated time series based on both spatial and temporal similarities. In the temporal perspective, the pattern of a time series is flexibly modeled as

a mixture of Gaussian processes, with a Dirichlet process (DP) prior over mixture components. In the spatial perspective, the spatial location is incorporated as a feature for clustering, like a time series being incorporated as a feature. Namely, we model the spatial distribution of each cluster as a DP Gaussian mixture density. For the proposed model, the number of clusters does not need to be specified in advance, but rather is automatically determined during the clustering procedure. Moreover, the spatial distribution of each cluster can be flexibly modeled with multiple modes, without determining the number of modes or specifying spatial neighborhood structures in advance. Variational inference is employed for the efficient posterior computation of the proposed model. We validate the proposed model using simulated and real-data examples.

Split Regularized Regression

Anthony-Alexander Christidis, Laks Lakshmanan, Ezequiel Smucler & Ruben Zamar

Abstract

We propose an approach for fitting linear regression models that splits the set of covariates into groups. The optimal split of the variables into groups and the regularized estimation of the regression coefficients are performed by minimizing an objective function that encourages sparsity within each group and diversity among them. The estimated coefficients are then pooled together to form the final fit. Our procedure works on top of a given penalized linear regression estimator (e.g., Lasso, elastic net) by fitting it to possibly overlapping groups of features, encouraging diversity among these groups to reduce the correlation of the corresponding predictions. For the case of two groups, elastic net penalty and orthogonal predictors, we give a closed form solution for the regression coefficients in each group. We establish the consistency of our method with the number of predictors possibly increasing with the sample size. An extensive simulation study and real-data applications show that in general the proposed method improves the prediction accuracy of the base estimator used in the procedure. Possible extensions to GLMs and other models are discussed. The supplemental material for this article, available online, contains the proofs of our theoretical results and the full results of our simulation study.

A Unified Approach to Sparse Tweedie Modeling of Multisource Insurance Claim Data

Simon Fontaine, Yi Yang, Wei Qian, Yuwen Gu & Bo Fan

Abstract

Actuarial practitioners now have access to multiple sources of insurance data corresponding to various situations: multiple business lines, umbrella coverage, multiple hazards, and so on. Despite the wide use and simple nature of single-target approaches, modeling these types of data may benefit from an approach performing variable selection jointly across the sources. We propose a unified algorithm to perform sparse learning of such fused insurance data under the Tweedie (compound Poisson) model. By integrating ideas from multitask sparse learning and sparse Tweedie modeling, our algorithm produces flexible regularization that balances predictor sparsity and between-sources sparsity. When applied to simulated and real data, our approach clearly outperforms single-target modeling in both prediction and selection accuracy, notably when the sources do not have exactly the same set of predictors. An efficient implementation of the proposed algorithm is provided in our R package MStweedie, which is available at https://github.com/fontaine618/MStweedie. Supplementary materials for this article are available online.

Model Misspecification of Generalized Gamma Distribution for Accelerated Lifetime-Censored Data

Marzieh Khakifirooz, Sheng Tsaing Tseng & Mahdi Fathi

Abstract

The performance of reliability inference strongly depends on the modeling of the product's lifetime distribution. Many products have complex lifetime distributions whose optimal settings are not easily found. Practitioners prefer to use simpler lifetime distribution to facilitate the data modeling process while knowing the true distribution. Therefore, the

effects of model mis-specification on the product's lifetime prediction is an interesting research area. This article presents some results on the behavior of the relative bias (RB) and relative variability (RV) of pth quantile of the accelerated lifetime (ALT) experiment when the generalized Gamma (GG_3) distribution is incorrectly specified as Lognormal or Weibull distribution. Both complete and censored ALT models are analyzed. At first, the analytical expressions for the expected log-likelihood function of the misspecified model with respect to the true model is derived. Consequently, the best parameter for the incorrect model is obtained directly via a numerical optimization to achieve a higher accuracy model than the wrong one for the end-goal task. The results demonstrate that the tail quantiles are significantly overestimated (underestimated) when data are wrongly fitted by Lognormal (Weibull) distribution. Moreover, the variability of the tail quantiles is significantly enlarged when the model is incorrectly specified as Lognormal or Weibull distribution. Precisely, the effect on the tail quantiles is more significant when the sample size and censoring ratio are not large enough. Supplementary materials for this article are available online.

Modification of the Maximin and ϕ_p (Phi) Criteria to Achieve Statistically Uniform Distribution of Sampling Points

Miroslav Vořechovský & Jan Eliáš

Abstract

This article proposes a sampling technique that delivers robust designs, that is, point sets selected from a design domain in the shape of a unit hypercube. The designs are guaranteed to provide a *statistically uniform* point distribution, meaning that every location has the same probability of being selected. Moreover, the designs are *sample uniform*, meaning that each individual design has its points spread evenly throughout the domain. The *sample uniformity* (often measured via a *discrepancy* criterion) is achieved using distance-based criteria (ϕp or Maximin), that is, criteria normally used in space-filling designs. We show that the standard intersite metrics employed in distance-based criteria (Maximin and ϕp (phi)) do *not* deliver statistically uniform designs. Similarly, designs optimized via centered L_2 discrepancy or support points are also not statistically uniform. When these designs (after optimization based on intersite distances) are used for Monte Carlo type of integration, their statistical nonuniformity is a serious problem as it may lead to a systematic bias. This article proposes using a periodic metric to guarantee the statistical uniformity of the family of distance-based designs. The presented designs used as benchmarks in the article are only taken from the class of Latin hypercube designs, which forces univariate projections to be uniform and improves accuracy in Monte Carlo integration of some functions. Supplementary materials for this article are available online.

Sliced Designs for Multi-Platform Online Experiments

Soheil Sadeghi, Peter Chien & Neeraj Arora

Abstract

Multivariate testing is a popular method to improve websites, mobile apps, and email campaigns. A unique aspect of testing in the online space is that it needs to be conducted across multiple platforms such as a desktop and a smartphone. The existing experimental design literature does not offer precise guidance for such a multi-platform context. In this article, we introduce a multi-platform design framework that allows us to measure the effect of the design factors for each platform and the interaction effect of the design factors with platforms. Substantively, the resulting designs are of great importance for testing digital campaigns across platforms. We illustrate this in an empirical email application to maximize engagement for a digital magazine. We introduce a novel "sliced effect hierarchy principle" and develop design criteria to generate factorial designs for multi-platform experiments. To help construct such designs, we prove a theorem that connects the proposed designs to the well-known minimum aberration designs. We find that experimental versions made for one platform should be similar to other platforms. From the standpoint of real-world application, such homogeneous subdesigns are cheaper to implement. To assist practitioners, we provide an algorithm to construct the designs that we propose.

Bradley Jones, Ryan Lekivetz, Dibyen Majumdar, Christopher J. Nachtsheim & Jonathan W. Stallrich

Abstract

In this article, we propose a new method for constructing supersaturated designs that is based on the Kronecker product of two carefully chosen matrices. The construction method leads to a partitioning of the factors of the design such that the factors within a group are correlated to the others within the same group, but are orthogonal to any factor in any other group. We refer to the resulting designs as *group-orthogonal supersaturated designs*. We leverage this group structure to obtain an unbiased estimate of the error variance, and to develop an effective, design-based model selection procedure. Simulation results show that the use of these designs, in conjunction with our model selection procedure enables the identification of larger numbers of active main effects than have previously been reported for supersaturated designs. The designs can also be used in group screening; however, unlike previous group-screening procedures, with our designs, main effects in a group are not confounded. Supplementary materials for this article are available online.

Ridge Regression: A Historical Context

Roger W. Hoerl

Abstract

Two classical articles on Ridge Regression by Arthur Hoerl and Robert Kennard were published in *Technometrics* in 1970, making 2020 their 50th anniversary. The theory and practice of Ridge Regression, and of related biased shrinkage estimators, have been extensively developed over the years. Further, newer shrinkage estimators, such as the Lasso and the Elastic Net, have become popular more recently. These newer developments have led to renewed interest in the original 1970 articles. What has perhaps been lost since 1970 is the context of these classic articles. That is, who were Art Hoerl and Bob Kennard, and what led two statisticians working in the private sector to develop Ridge Regression in the first place? What are the origins of Ridge Regression? Where did the name come from? The purpose of this article is to provide this historical context by discussing the men involved, their work at DuPont, and their approach to methodological development. As Art Hoerl was my father, this is admittedly a personal viewpoint.

Ridge Regularization: An Essential Concept in Data Science

Trevor Hastie

Abstract

Ridge or more formally l2 regularization shows up in many areas of statistics and machine learning. It is one of those essential devices that any good data scientist needs to master for their craft. In this brief *ridge fest*, I have collected together some of the magic and beauty of ridge that my colleagues and I have encountered over the past 40 years in applied statistics.

Generalized Principal Component Analysis: Projection of Saturated Model Parameters

Andrew J. Landgraf & Yoonkyung Lee

Abstract

Principal component analysis (PCA) is very useful for a wide variety of data analysis tasks, but its implicit connection to the Gaussian distribution can be undesirable for discrete data such as binary and multi-category responses or counts. We generalize PCA to handle various types of data using the generalized linear model framework. In contrast to the existing approach of matrix factorizations for exponential family data, our generalized PCA provides low-rank estimates of the natural parameters by projecting the saturated model parameters. This difference in formulation leads to the favorable properties that the number of parameters does not grow with the sample size and simple matrix multiplication suffices for computation of the principal component scores on new data. A practical algorithm which can incorporate missing data and case weights is developed for finding the projection matrix. Examples on simulated and real count data show the improvement of generalized PCA over standard PCA for matrix completion, visualization, and collaborative filtering.

Sparse Principal Component Analysis Based on Least Trimmed Squares

Yixin Wang & Stefan Van Aelst

Abstract

Sparse principal component analysis (PCA) is used to obtain stable and interpretable principal components (PCs) from high-dimensional data. A robust sparse PCA method is proposed to handle potential outliers in the data. The proposed method is based on the least trimmed squares PCA method which provides robust but non-sparse PC estimates. To obtain sparse solutions, our method incorporates a regularization penalty on the loading vectors. The principal directions are determined sequentially to avoid that outliers in the PC subspace destroy the sparse structure of the loadings. Simulation studies and real data examples show that the new method gives accurate estimates, even when the data are highly contaminated. Moreover, compared to existing robust sparse PCA methods the computation time is reduced to a great extent. Supplementary materials providing more simulation results and discussion, and an R package to compute the proposed method are available online.

Sequential Model-Based Optimization for Continuous Inputs with Finite Decision Space

Qiong Zhang & Youngdeok Hwang

Abstract

Optimization using stochastic computer experiments is commonplace in engineering and industry. This article addresses the problem of optimization, in which the input space of stochastic computer model is continuous, whereas the decision space in the real problem is restricted to be finite. We propose a sequential model-based optimization method to tackle this problem. The knowledge gradient based on this restricted decision space is used as the criterion to choose new design points, and the fixed rank kriging or Gaussian process is used as the surrogate. This combination takes advantage of the continuous search space to decrease the uncertainty over the finite decision space. We demonstrate the benefit of our proposed methodology compared with existing sequential model-based optimization methods.

Adaptive Minimum Confidence Region Rule for Multivariate Initialization Bias Truncation in Discrete-Event Simulations

Jianguo Wu, Honglun Xu, Feng Ju & Tzu-Liang (Bill) Tseng

Abstract

Initialization bias truncation is critically important for system performance assessment and warm-up length estimation in discrete-event simulations. Most of the existing methods are for univariate signals, while multivariate truncation has been rarely studied. To fill such gap, this article proposes an efficient method, called adaptive minimum confidence region rule (AMCR) for multivariate initialization bias truncation. It determines the truncation point by minimizing the modified confidence volume with a tuning parameter for the mean estimate. An elbow method is developed for adaptive selection of the tuning parameter. Theoretical properties of the AMCR rule for both data with and without autocorrelations have been derived for justification and practical guidance. The effectiveness and superiority of the AMCR rule over other existing approaches have been demonstrated through thorough numerical studies and real application.

Uncertainty Quantification for High-Dimensional Sparse Nonparametric Additive Models

Qi Gao, Randy C. S. Lai, Thomas C. M. Lee & Yao Li

Abstract

Statistical inference in high-dimensional settings has recently attracted enormous attention within the literature. However, most published work focuses on the parametric linear regression problem. This article considers an important extension of this problem: statistical inference for high-dimensional sparse nonparametric additive models.

To be more precise, this article develops a methodology for constructing a probability density function on the set of all candidate models. This methodology can also be applied to construct confidence intervals for various quantities of interest (such as noise variance) and confidence bands for the additive functions. This methodology is derived using a generalized fiducial inference framework. It is shown that results produced by the proposed methodology enjoy correct asymptotic frequentist properties. Empirical results obtained from numerical experimentation verify this theoretical claim. Lastly, the methodology is applied to a gene expression dataset and discovers new findings for which most existing methods based on parametric linear modeling failed to observe.

Transformation and Additivity in Gaussian Processes

Li-Hsiang Lin & V. Roshan Joseph

Abstract

We discuss the problem of approximating a deterministic function using Gaussian processes (GPs). The role of transformation in GP modeling is not well understood. We argue that transformation of the response can be used for making the deterministic function approximately additive, which can then be easily estimated using an additive GP. We call such a GP a transformed additive Gaussian (TAG) process. To capture possible interactions which are unaccounted for in an additive model, we propose an extension of the TAG process called transformed approximately additive Gaussian (TAAG) process. We develop efficient techniques for fitting a TAAG process. In fact, we show that it can be fitted to high-dimensional data much more efficiently than a standard GP. Furthermore, we show that the use of the TAAG process leads to better estimation, interpretation, visualization, and prediction. The proposed methods are implemented in the R package *TAG*.

Statistical Modeling of Multivariate Destructive Degradation Tests With Blocking

Qiuzhuang Sun, Zhi-Sheng Ye & Yili Hong

Abstract

In degradation tests, the test units are usually divided into several groups, with each group tested simultaneously in a test rig. Each rig constitutes a rig-layer block from the perspective of design of experiments. Within each rig, the test units measured at the same time further form a gauge-layer block. Due to the uncontrollable factors among test rigs and the common errors incurred for each measurement, the degradation measurements of the test units may differ among various blocks. On the other hand, the degradation should be more homogeneous within a block. Motivated by an application of emerging contaminants (ECs), this study proposes a multivariate statistical model to account for the two-layer block effects in destructive degradation tests. A multivariate Wiener process is first used to model the correlation among different dimensions of degradation. The rig-layer block effect is modeled by a one-dimensional frailty motivated by the degradation physics, while the gauge-layer block effect at each measurement epoch is captured by a common additive measurement error. We develop an expectation-maximization algorithm to obtain the point estimates of the model parameters and construct confidence intervals for the parameters. A procedure is proposed to test significance of the block effects in the degradation data. Through a case study on an EC degradation dataset, we show the existence of the two-layer block effects from the test. By making use of the proposed model, decision makers can readily make risk assessment of each contaminant and determine the minimal water treatment time for removal of the contaminants.

A Note on Cross-Validation for Lasso Under Measurement Errors

Abhirup Datta & Hui Zou

Abstract

Variants of the Lasso or ℓ_1-penalized regression have been proposed to accommodate for presence of measurement errors in the covariates. Theoretical guarantees of these estimates have been established for some oracle values of the regularization parameters which are not known in practice. Data-driven tuning such as cross-validation has not been studied when covariates contain measurement errors. We demonstrate that in the presence of error-in-covariates, even

when using a Lasso-variant that adjusts for measurement error, application of naive leave-one-out cross-validation to select the tuning parameter can be problematic. We provide an example where such a practice leads to estimation inconsistency. We also prove that a simple correction to cross-validation procedure restores consistency. We also study the risk consistency of the two cross-validation procedures and offer guideline on the choice of cross-validation based on the measurement error distributions of the training and the prediction data. The theoretical findings are validated using simulated data.