Centrum für
Europa-,
Governance- und
Entwicklungsforschung

cege-Report
April 2013

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Verteilungsregression: Statistische Modellierung jenseits des Mittelwertes

Vorurteile zur Statistik beziehen sich gerne auf die ausschließliche Betrachtung von Mittelwerten als Repräsentanten von Grundgesamtheiten. Dies drückt sich insbesondere in der häufig verwendeten Formulierung aus, dass bestimmte Eigenschaften „statistisch gesehen" zutreffen:

- Statistisch gesehen gebiert jede Frau in Deutschland 1,3 Kinder in ihrem Leben.
- Statistisch gesehen besitzt jeder Deutsche etwa 88.000 Euro Sparguthaben.
- Steht man mit einem Fuß auf einer heißen Herdplatte und mit dem anderen Fuß in der Gefriertruhe, so fühlt man sich statistisch gesehen wohl.

Diese Beispiele werden dann gerne herangezogen, um die Unsinnigkeit statistischer Analysen zu illustrieren und die Behauptung aufzustellen, Statistik habe mit dem wirklichen Leben keine Verbindung.

Um über die reine Betrachtung von Mittelwerten hinaus Informationen über die Verteilung von Merkmalen in Grundgesamtheiten zu ermitteln, ist es in der wissenschaftlichen Analyse mittlerweile üblich, neben dem Mittelwert auch Maßzahlen für beispielsweise die Streuung oder Schiefe der Verteilung anzugeben oder direkt grafische Darstellungen zu verwenden. Damit wird ein deutlich facettenreicheres und detaillierteres Bild vermittelt, aus dem sich im obigen Beispiel der Sparguthaben beispielsweise ergibt, dass die Verteilung der Sparguthaben extrem rechtsschief ist, so dass eine Konzentration auf relativ geringen Beträgen mit einer relativ geringen Zahl sehr großer Guthaben einhergeht.

Leider hat sich diese allgemeinere Betrachtung von Verteilungen bisher im Wesentlichen nur für die Beschreibung einzelner Merkmale und nicht in der statistischen Modellierung von Zusammenhängen durchgesetzt. In Bereich der statistischen Modellierung findet man weiterhin einen sehr starken Fokus auf der Beschreibung erwarteter oder durchschnittlicher Zusammenhänge.

Interessiert man sich beispielsweise für den Zusammenhang zwischen den einer Firma zur Verfügung stehenden Ressourcen (Input) und dem erzeugten Umsatz (Output), so interessiert üblicherweise weniger, welchen Output eine durchschnittliche Firma erzielt, sondern welchen Output eine überdurchschnittlich erfolgreiche Firma im Vergleich mit durchschnittlichen Firmen (bei gleichem Input) erreichen kann. Dies führt zur Schätzung sogenannter technischer Effizienzen, die die überdurchschnittliche Performance von Firmen quantifizieren. Die Bestimmung technischer Effizienzen beruht dann beispielsweise auf der Schätzung von Produktionsfunktionen bzw. Frontiers, also oberen Produktionsgrenzen, die mit gegebenen Mitteln erreicht werden können.

Auch jenseits dieses Beispiels ist in den Wirtschaftswissenschaften die Beschreibung allgemeinerer Zusammenhänge als nur zwischen Einflussgrößen und Mittelwert einer Zielgröße von wesentlichem Interesse. So beschäftigt sich die Analyse der Ungleichheit explizit mit der Verteilung von Gütern oder Ressourcen und nicht oder nicht nur mit der durchschnittlich zur Verfügung stehenden Menge an Ressourcen. Insbesondere interessiert die Bestimmung von Determinanten der Einkommensungleichheit oder der Entwicklung von Armutsrisiken.

Die Statistik entwickelt hierzu moderne Modellierungsansätze, die sich unter dem Schlagwort der Verteilungsregression zusammenfassen lassen, Diese zielen auf die Beschreibung der gesamten Verteilung einer Zielgröße (wie des in der Produktion erreichten Outputs) in Abhängigkeit von Einflussgrößen (Inputs) ab. Die Modellklasse beinhaltet insbesondere die Entwicklung von Schätzverfahren für flexiblere Verteilungsmodelle aber auch von verteilungsfreien Ansätzen wie der Quantilregression. Damit verbunden ist die Umsetzung in entsprechender Software, um die Verbreitung der entwickelten Verfahren sicherzustellen.

Als Fazit bleibt festzuhalten, dass Statistik nicht mit der Betrachtung von Mittelwerten gleichzusetzen ist und die statistische Modellierung jenseits des Mittelwerts zum Handwerkszeug quantitativ arbeitender Wissenschaftler gehören sollte.

Prof. Dr. Thomas Kneib
Statistik, *cege*

Optimal Ageing and Death

Prof. Dr. Holger Strulik

In this research project, funded under the EU FP7 program, we integrated a physiologically founded aging process of the human body into an otherwise fairly standard economic life cycle model. Broadly speaking, the representative households find investments in health attractive as these expenditures serve to slow down the aging process through maintenance and repair, allowing for greater life span and thus higher welfare. At the same time, households value other things that do not prolong life. These expenditures naturally constitute the alternative costs of health spending. Within this setting the first key question is how households would optimally want to allocate their income over the life cycle. Once this choice has been made, the model predicts the (optimal) aging and timing of household death.

The fact that the underlying model parameters have a foundation in the natural sciences allowed for a rigorous calibration of the model, which we did by matching an average 35 years old US male in the year 2000. The model was then used to make out of sample predictions for health behavior and life expectancy across the world.

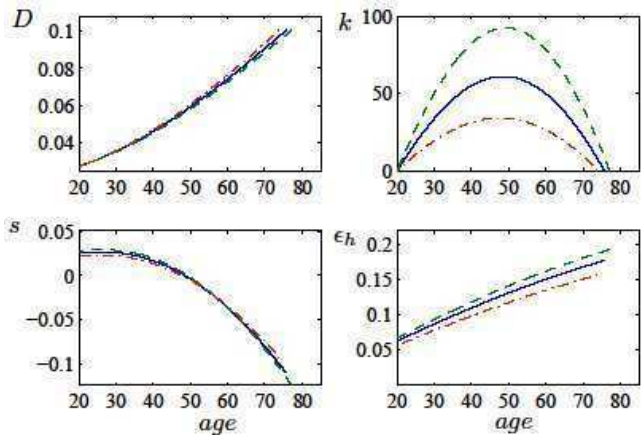
Figure 1 shows how the average male US American is predicted to react if his income is perturbed. The blue line shows life-trajectories for health deficit accumulation (top left panel), capital accumulation (top right panel), the savings rate (bottom left), and the health expenditure share (bottom right) from the age of 20 to death. The green (dashed) line is associated with an increase of labor income by 1/3, the red (dashed-dotted) line depicts the reaction to a reduction of labor income by 1/3. As can be seen from the figure, the consequence of higher income is an increase in longevity, peak wealth, and the share of health spending.

The intuition for why higher income leads to longer life is found in the desire to smooth consumption. With increasing income, health spending rises more than regular consumption. This occurs as the incentive to smooth the latter is relatively stronger due to diminishing per period marginal utility. Higher income therefore leads to a larger adjustment in the level of health spending compared to non-health consumption. The end result is a slower speed of aging, and a longer life.

The issue of main interest, however, is the quantitative impact on longevity. As seen from the top left hand side corner of the figure, the impact is modest though not inconsequential. An increase of income of 1/3 (achievable in a generation with an income growth rate of about 1% per year) translates into an increase in longevity of 1.7 years; the reduction involves a fall in longevity of 1.6 years. If we convert the impact into an elasticity - the elasticity of longevity with respect to income - we find it to be around 0.08, implying that a 100 percent increase of income would result in an 8 percent increase of life-expectancy.

The next experiment concerned health costs; a doubling of the relative price of health goods. Rising relative health prices is a realistic scenario since the price index of medical care has risen faster than the price index of GDP. When the relative price of health increases, individuals substitute towards regular consumption. As a result, the health share declines. With less health investments, savings decline as well. The end result of a doubling of the relative health price for longevity is a reduction by 2.5 years. This amounts to a longevity-price elasticity of 0.09. While this elasticity is of roughly the same numerical size as the income elasticity it is well to bear in mind that the potential for variation is far smaller; income levels vary across the world easily by a factor of 30 or more, far towering any cross-country variation in price levels.

We also investigated the impact of health productivity, i.e. generally rising medical effectiveness. If health productivity rises by 1/3 the consequence is an increase in life expectancy at 20 of nearly a decade. The implied elasticity is about 1/2. This is a very large effect, suggesting that the impact from improvements in health productivity easily may have towered that of rising income. It also suggests that improvements in health efficiency have a large potential to improve life expectancy.



The model does remarkably well in accounting for the income gradient in cross country data (the so-called Preston curve). A first key result is that prevailing differences in average income across countries (and thus in implied health investments) can account for a difference in life expectancy at age 20 of about a decade. Income therefore matters a great deal. However, a second key result of the analysis is that technological change (rising in health efficiency) has a much greater impact on longevity; the impact from technology is also much larger than that of changes in health services. From a policy perspective this suggests that improvements in health efficiency, which might involve institutional reform of the health sector, will likely have a much bigger impact on health of the populations than policies geared towards subsidizing health spending by individuals.

Prof. Dr. Holger Strulik, *cege*

The Effect of Improved Sanitation on Diarrheal Morbidity of Children in India

Reducing child mortality by two-thirds between 1990 and 2015 is one of the key goals of the United Nations Millennium Development program. Two regions, South Asia and Sub-Saharan Africa, share the greatest burden of child mortality. Globally, seven million children under the age of five die every year and about 11% of these deaths are caused by diarrhea (UNICEF, 2012). Diarrhea is considered to be the second leading cause of under-five mortality after pneumonia. Among all countries severely affected by diarrheal mortality, India alone accounts for 0.2 million diarrheal deaths per year out of 0.7 million diarrheal deaths worldwide in 2011. Though many factors contribute to diarrheal deaths, one of the key factors contributing to the frequency and burden of diarrheal disease is the pronounced lack of water and sanitation in a majority of developing countries (Zwane and Kremer, World Bank Research Observer 2007).

In addition to unsafe drinking water and unhygienic behavior, lack of basic sanitation significantly contributes to the diarrheal disease burden. Open defecation, which is still practiced by around 1.1 billion people worldwide, remains a major contributing factor to diarrheal disease. As per national figures from the Government of India (GOI), 54% of India's population openly defecates and it was anguishly termed as "World's Capital for Open Defecation" by Drinking Water and Sanitation Minister Jairam Ramesh as India tops the global list in open defecation.

Evaluation hypothesis, data, and methodology

The GOI has invested a large amount of funds in the Total Sanitation Campaign (TSC), it is thus important to evaluate if and to what extent these efforts contributed to reducing the burden of diarrheal disease. TSC started in 1999 and continued until 2009, $1.6 billion have been spent under this program to build 57 million household toilets. Unfortunately, in the absence of a randomized evaluation of this sanitation program, it is very challenging to isolate the causal effect of sanitation on health. In a recently published paper, we used the method of propensity score matching (PSM) to estimate the effect of access to improved sanitation on children's probability of having diarrhea (Kumar and Vollmer, Health Economics 2012). The basic idea in PSM is to mimic treatment and control groups through matching households, which have similar observed characteristics such that comparisons can be made between these matched groups. We used the third wave of the District Level Household Survey (DLHS) for our analysis.

Findings

Evidence on the causal effect of sanitation on health outcomes in India is quite weak and inconclusive. Duraiswamy (National Council of Applied Economic Research Working paper 2001) found no significant effect of availability of a toilet, hand-washing behavior, and sources of drinking water on children's morbidity. Fan and Mahal (Journal of Development Effectiveness 2011) apply matching methods to estimate the effects of water supply, toilets and hand washing on child diarrhea in rural India. They find no significant effect of improved sanitation on diarrheal morbidity. In contrast, our study shows that children in households with access to improved sanitation have a lower diarrheal incidence than children in households without. The incidence of diarrhea for children living in a household with improved sanitation is 2.2 percentage points lower than for children living in a household without improved sanitation. Put differently, the odds of contracting diarrhea, for children in households without improved sanitation, are 24% higher than for children in households with improved sanitation.

We also find that the effect of sanitation varies by socioeconomic status (SES) and other household characteristics. We find a statistically significant treatment effect of 2.5 percentage points in the highest SES group, whereas the effect in the middle and low SES groups is statistically insignificant. We find a similar result when we stratify by using a variable that can be considered as a proxy for health and hygiene behavior: Whether or not a household treats water before drinking it. We find a statistically significant treatment effect of 3.3 percentage points for households that treat their water before drinking, whereas the effect is insignificant for households that do not treat their water. We also find that the treatment effect is only significant for boys and not for girls.

Conclusions

India made huge progress in sanitation coverage in rural areas with the launch of TSC. However, it still has the highest burden of child mortality and morbidity related to diarrhea in South Asia. Our paper indicates positive health benefits of access to improved sanitation, but it also highlights that the benefits have not been fully realized yet. Our results also underscore the importance of hygiene behavior, such as boiling drinking water and washing hands with soap. To fully eradicate the prevalence of diarrheal deaths in India, a complementation of TSC with Water Sanitation & Hygiene (WASH) programs is the need of the hour. Continuing improvements in sanitation infrastructure with complementary policies for behavior change through community participation, education, awareness, and health promotion activities may go a long way in reducing diarrheal burden in India.

Santosh Kumar, Institute for Health Metrics and Evaluation, University of Washington

SebastianVollmer, Universität Göttingen, *cege*

Aktuelle Publikationen in referierten Journals

- **K. Bizer**/J. Thomä (2013): To protect or not to protect? Modes of appropriability in the small enterprise sector., Research Policy 42(1).
- **A. Gehringer** (2013), Financial Liberalization, Growth, Productivity and Capital Accumulation: The Case of European Integration, in: International Review of Economics & Finance, 25.
- **H. Herwartz**/F. Siedenburg (2013): To converge or not converge: unit labor cost inflation in the Euro area, Empirical Economics 44.
- **A. Oestreicher**/R. Koch/D. Vorndamme (2012), Reforming Inter-period Loss-offset Provisions, in: Die Betriebswirtschaft/Business Administration Review (BARev), 72.
- O. Himmler/**R. Schwager** (2013): Double standards in educational standards: Do schools with a disadvantaged student body grade more leniently?, German Economic Review 14.
- **S. Vollmer**/H. Holzmann/F. Ketterer/**S. Klasen** (2013), Distribution Dynamics of Regional GDP per Employee, in: Unified Germany. Empirical Economics 44(2).
- **S. Vollmer**/H. Holzmann/F. Ketterer/**S. Klasen**/D. Canning (2013), The Emergence of Three Human Development Clubs. PLOS ONE 8(3).
- S. Kumar/**S. Vollmer** (2013) Does Access to Improved Sanitation Reduce Childhood Diarrhea in Rural India? Health Economics 22(4).

cege-Diskussionspapiere

Auf der *cege*-Homepage stehen neue *cege*-Diskussionspapiere als Download zur Verfügung.

- No. 152: **Strulik, Holger** / Werner, Katharina: „50 is the New 30 – Long-run Trends of Schooling and Retirement Explained by Human Aging."
- No. 151: Dalgaard, Carl-Johan / **Strulik, Holger**: „The History Augmented Solow Model."
- No. 150: **Strulik, Holger** / Trimborn, Timo: „The Dark Side of Fiscal Stimulus."
- No. 149: Prettner, Klaus: „Public education, technological change and economic prosperity."
- No. 148: Lankau, Matthias / Bicskei, Marianna / **Bizer, Kilian**: „Cooperation Preferences in the Provision of Public Goods: An Experimental Study on the Effects of Social Identity."
- No. 147: Krenz, Astrid: „Modeling Services Sectors' Agglomeration within a New Economic Geography Model."
- No. 146: Krenz, Astrid: „A Panel Co-integration Analysis of Industrial and Services Sectors' Agglomeration in the European Union."

cege-Forschungskolloquium im Sommersemester 2013

- 17. April 2013: Prof. Dr. Zohal Hessami (Universität Konstanz) „Academic Performance and Single-Sex Schooling: Evidence from a Natural Experiment in Switzerland"
- Sondertermin: Dienstag, 23. April 2013, 18:00 - 19:30 Uhr, Raum Oec 1.163: Prof. Ejan Mackaay, LL.D. (Université de Montreal) „If L&E is about impacts, what's its place in civil law reasoning?"
- 24. April 2013: Prof. Dr. Ulrike Leopold-Wildburger (Universität Graz) „Experimental Studies on Managerial Reporting"
- 8. Mai 2013: Prof. Dr. Dorothea Kübler (Wissenschaftszentrum Berlin) „Privacy Concerns and Unraveling in a Lemons Market: An Experiment"
- 15. Mai 2013: Prof. Dr. Christian Zehnder (Université Lausanne) „Delegating Responsibility to the Market: How Competition Alters Fairness Perceptions"
- 22. Mai 2013: Prof. Dr. Annika Herr (Universität Düsseldorf) „Pharmaceutical Prices under Regulation: Tiered Copayments and Reference Pricing in Germany"
- 12. Juni 2013: Prof. Dr. Volker Grossmann (Université Fribourg) „Migration, Capital Formation, and House Prices"
- 3. Juli 2013: Prof. Dr. Uwe Sunde (Universität München) „Geography, Epidemic Diseases, and Civil Conflict"
- 10. Juli 2013: Prof. Jennifer Jennings, Ph.D. (New York University) „The Summer Drug Holiday: Medical Adaptation to Academic Pressure and Psychiatric Financial Incentives"

Gastwissenschaftler im Sommersemester 2013

Gast an der Professur für Ökonometrie: Prof. Bernd Theilen (Universität Rovira I Virgili)

Impressum
Herausgeber: Centrum für Europa-, Governance- und Entwicklungsforschung (*cege*), Georg-August-Universität Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, **E-Mail:** cege@uni-goettingen.de, **Tel.** 0551 / 39 46 02, **Fax** 0551 / 39 19 55 8, **Internet:** http://www.cege.uni-goettingen.de, **Geschäftsführender Direktor:** Prof. Dr. Kilian Bizer, **Redaktion / Layout:** Prof. Dr. Kilian Bizer, Laura Birg, **Druck:** GWDG, Göttingen