ECONOMETRIC MODELING OF HEALTH CARE COSTS

John Mullahy
Univ. of Wisconsin-Madison

AHRQ/NCI Conference on
Health Care Costs: Standardized Methods and Estimates
for Research and Policy Applications

December 6-7, 2007

THE UNIVERSITY of WISCONSIN MADISON

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

Prescription Drug Event and Risk Adjustment Data

Prescription Drug Event

The Prescription Drug Event Front End System (PDFS) performs the initial file processing of the Prescription Drug Event data submitted by the Part D Plans. Upon completion of the intial file processing, the Prescription Drug Event data is then sent to the Drug Data Processing System for validation and authentication of the Medicare payment of covered drugs made by the Part D plans for their enrolled Medicare beneficiaries. Next, an analytic component, the Drug Benefit Calculator (DBC), loads the data into the National Medicare Drug Benefit Database and aggregates the data into data marts, when appropriate, to support drug, beneficiary, and plan analysis of incurred payments and payment reconciliation.

Risk Adjustment

The Front End Risk Adjustment System (FERAS) performs the initial file processing for risk adjustment data submitted by MA and MA-PD plans. Upon completion of the initial file processing, the risk adjustment data is sent to the Risk Adjustment Processing System (RAPS) at CMS for use in the calculation of Part C and Part D risk scores. These beneficiary level risk scores are passed to the Medicare Advantage Rx (MARx) system for use in Part C and Part D payment calculations.

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants

Ciaran S. Phibbs, Ph.D., Laurence C. Baker, Ph.D., Aaron B. Caughey, M.D., Ph.D., Beate Danielsen, Ph.D., Susan K. Schmitt, Ph.D., and Roderic H. Phibbs, M.D.

Health Technology Assessment

The Health Technology Assessment Program (HTA) provides high-quality information about the clinical effectiveness, cost-effectiveness, and broader impact of drugs, medical technologies, and health systems. Our impartial, rigorous, and comprehensive assessments examine four questions:

- How will this health technology affect the health of Canadians?
- How does it compare with alternatives?
- Does it provide value for the investment?
- Are there other health service implications to consider?

THE UNIVERSITY of WISCONSIN MADISON

CONGRESS OF THE UNITED STATES
CONGRESSIONAL BUDGET OFFICE

A
CBO
PAPER

MAY 2005

High-Cost Medicare Beneficiaries

Physician Profiling, Guiding Principles

See also:
Performance Measures Criteria
Pay-for-Performance

Preamble

The AAFP believes physician profiling is an analytic tool that uses epidemiological methods to compare physician practice patterns across various quality of care dimensions (process and clinical outcomes). Cost, service and resource utilization data are dimensions of measuring quality, but should not be used as independent measures of defining quality care. The ultimate goal is to improve clinical outcomes.

It is important to recognize that physician profiling is not intended to be used to address issues of physician competency, including the dimensions of medical knowledge, skills competence and physician performance. It is our belief that these issues should be addressed by the appropriate public and private credentialing bodies that exist for these purposes.

HEALTH ECONOMICS
Health Econ. (2007)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hec.1307

PROJECTIONS OF THE COSTS ASSOCIATED WITH COLORECTAL CANCER CARE IN THE UNITED STATES, 2000–2020†

K. ROBIN YABROFF*, ANGELA B. MARIOTTO, ERIC FEUER and MARTIN L. BROWN

Health Services and Economics Branch/Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

SUMMARY

Because of aging trends in the US, the number of prevalent colorectal cancer patients is expected to increase. We projected economic burden to the Medicare program and its beneficiaries through the year 2020. Burden was estimated for the initial phase of care, the period following diagnosis, the last year of life, and the continuing phase. Projected burden was evaluated with varying assumptions about incidence, survival, and costs of care. Estimated costs of care in 2000 in the initial, continuing, and last year of life phases of care were approximately $3.18 billion, $1.68 billion, and $2.63 billion, respectively. By the year 2020 under the 'fixed' current incidence, survival, and cost scenario, projected costs for the initial, continuing, and last year of life phases were $4.75 billion, $2.63 billion, and $4.05 billion. Under the current trends scenario (decreasing incidence, improving survival, and increasing costs), costs were $5.19 billion, $3.57 billion, and $5.27 billion. By the year 2020, estimated costs of colorectal cancer care among individuals aged 65 and older increased by 53% in the fixed scenario and by 89% in the current trends scenario. The future economic burden of colorectal cancer to the Medicare program and its beneficiaries in the US will be substantial. Published in 2007 by John Wiley & Sons, Ltd.

Received 30 August 2006; Revised 13 August 2007; Accepted 22 August 2007

KEY WORDS: colorectal neoplasms; cost of illness; costs and cost analysis; SEER Program; Medicare; prevalence

ROADMAP

1. How is empirical modeling useful in practice?

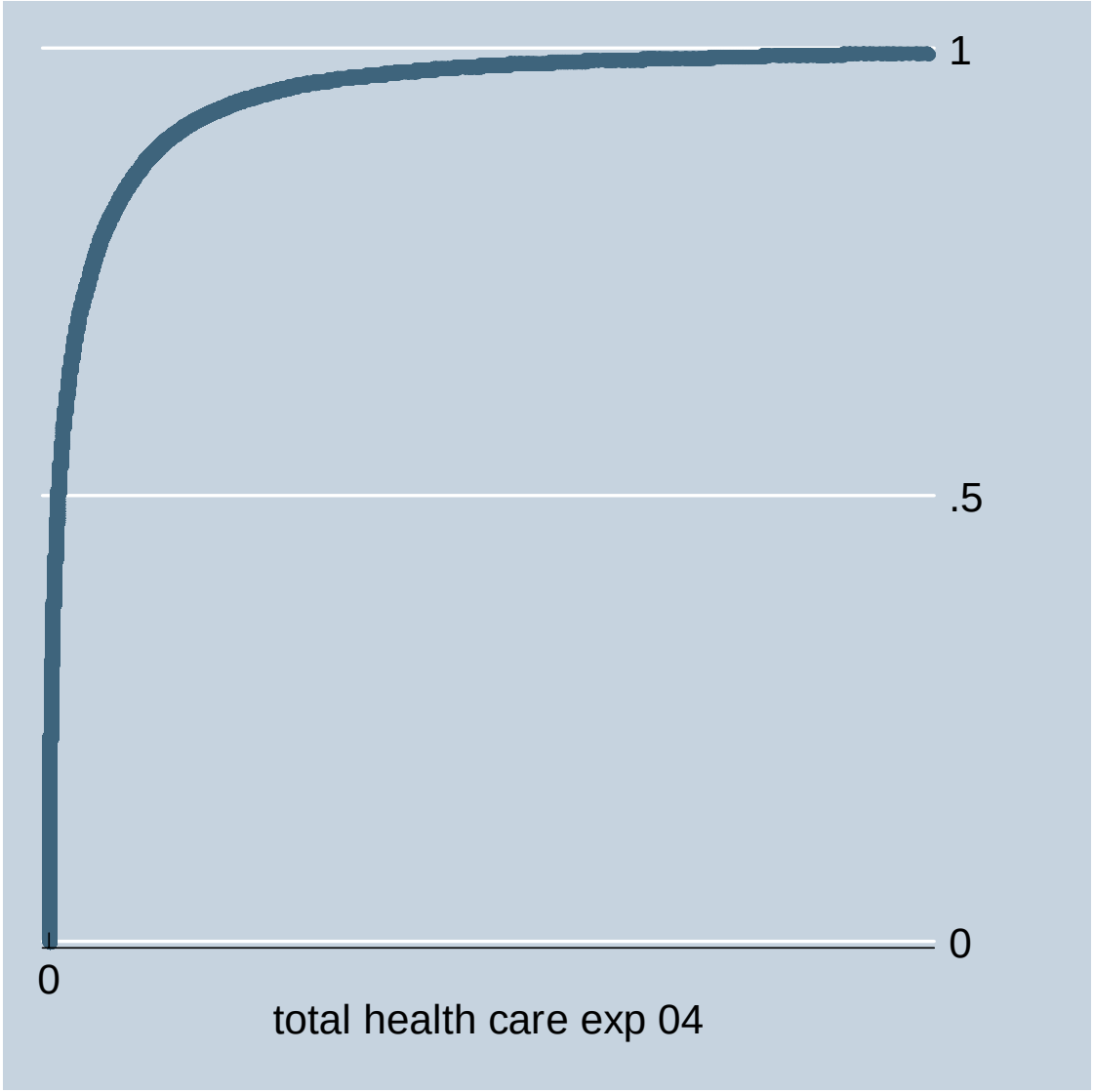
2. What sorts of questions may usefully be informed by such empirical investigations?

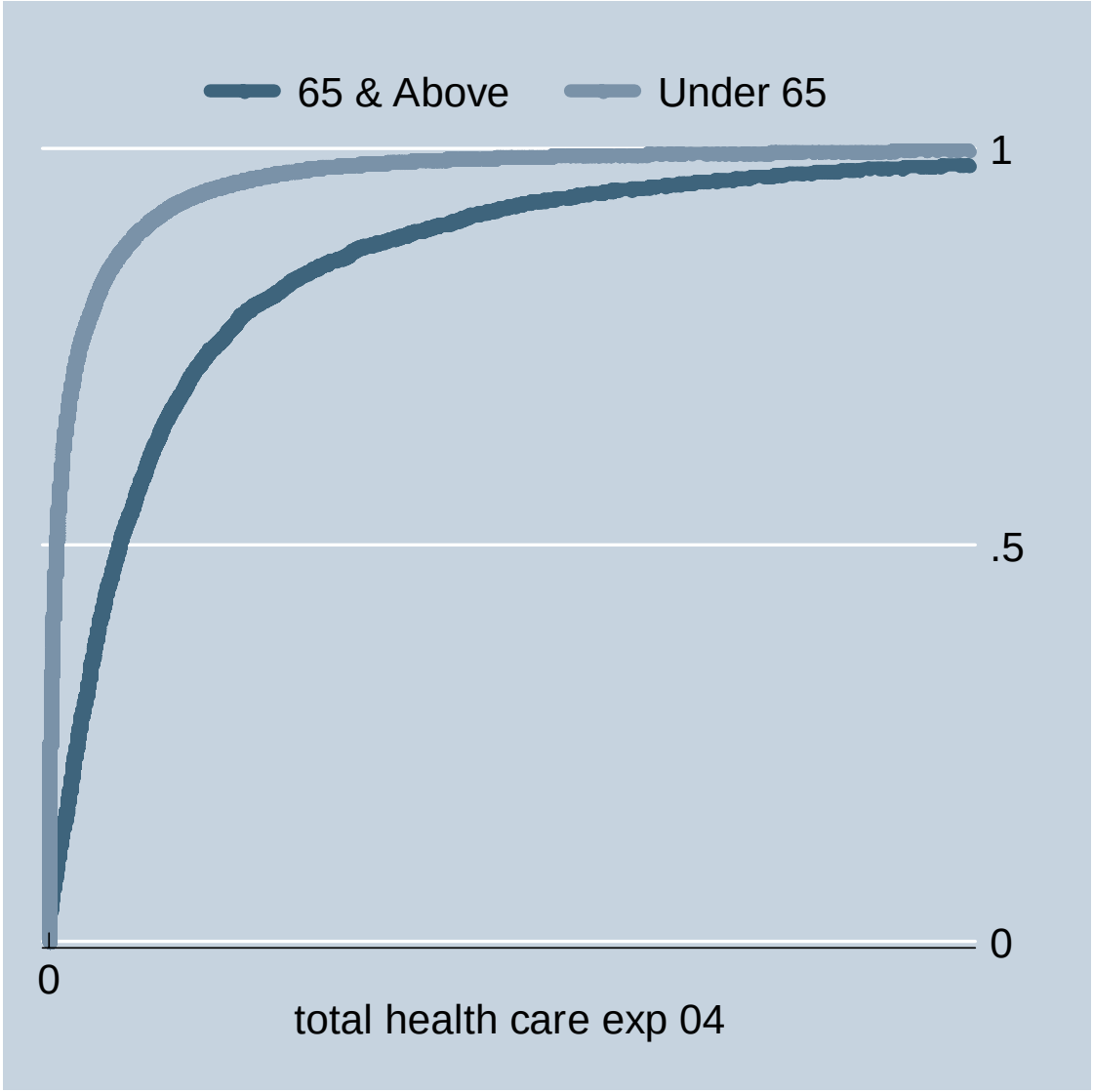
3. Specification and estimation of statistical models, in light of (2) and of the data

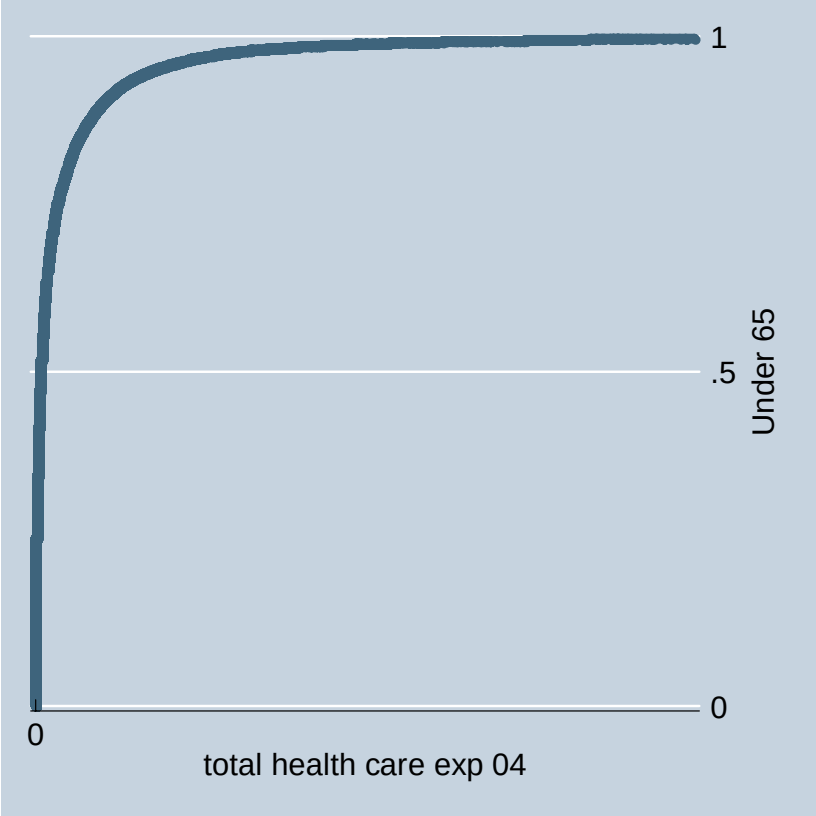
4. Goodness-of-Fit: Why and how?

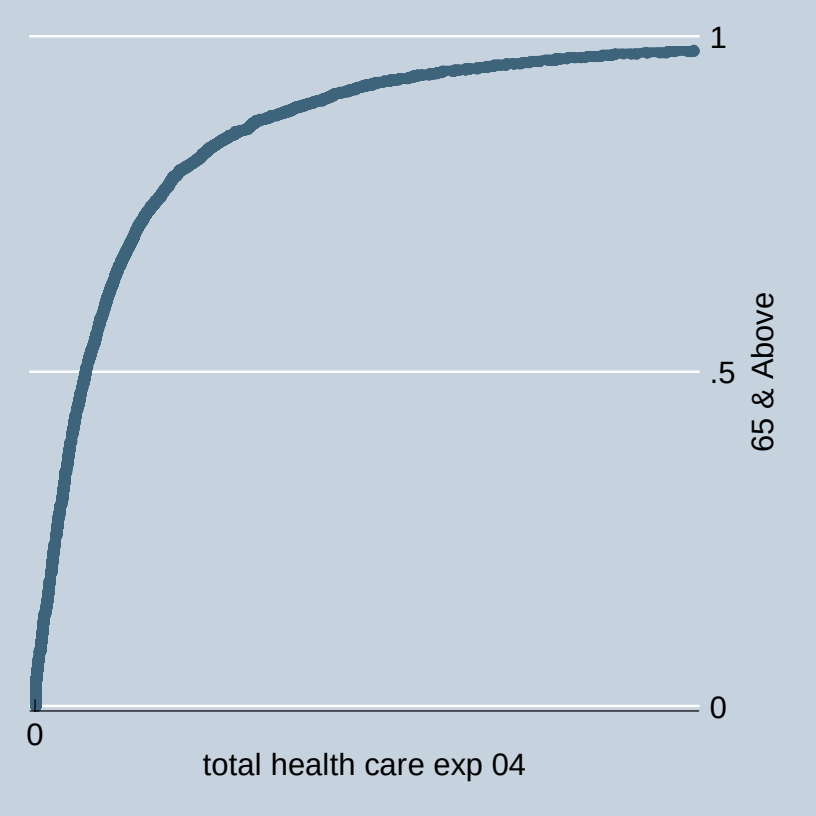
5. Considerations of "high-end" costs

6. Translating and reporting research usefully

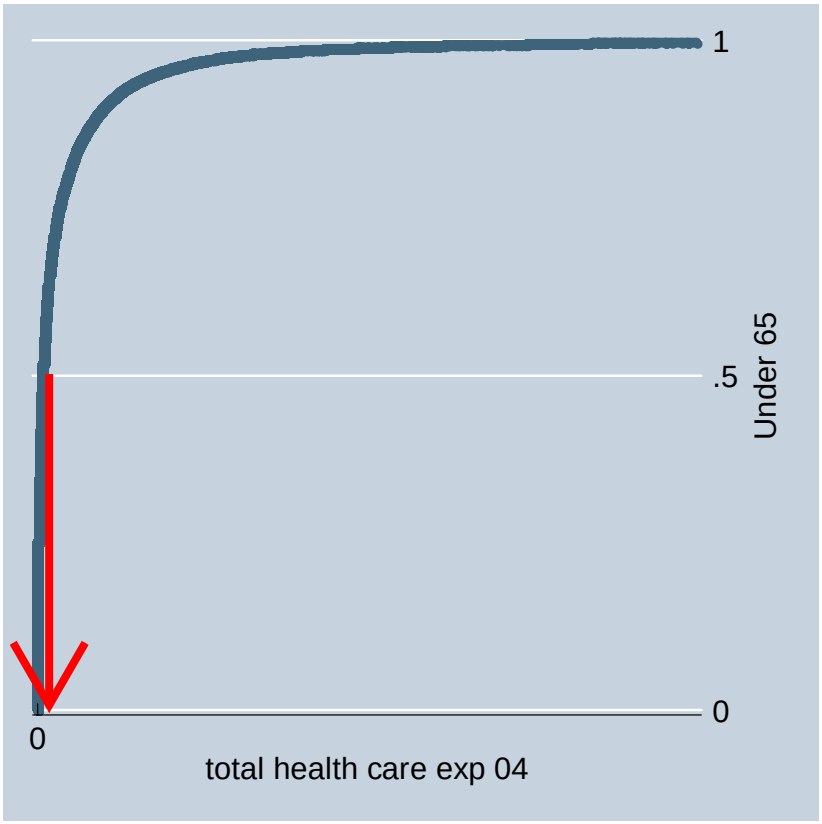


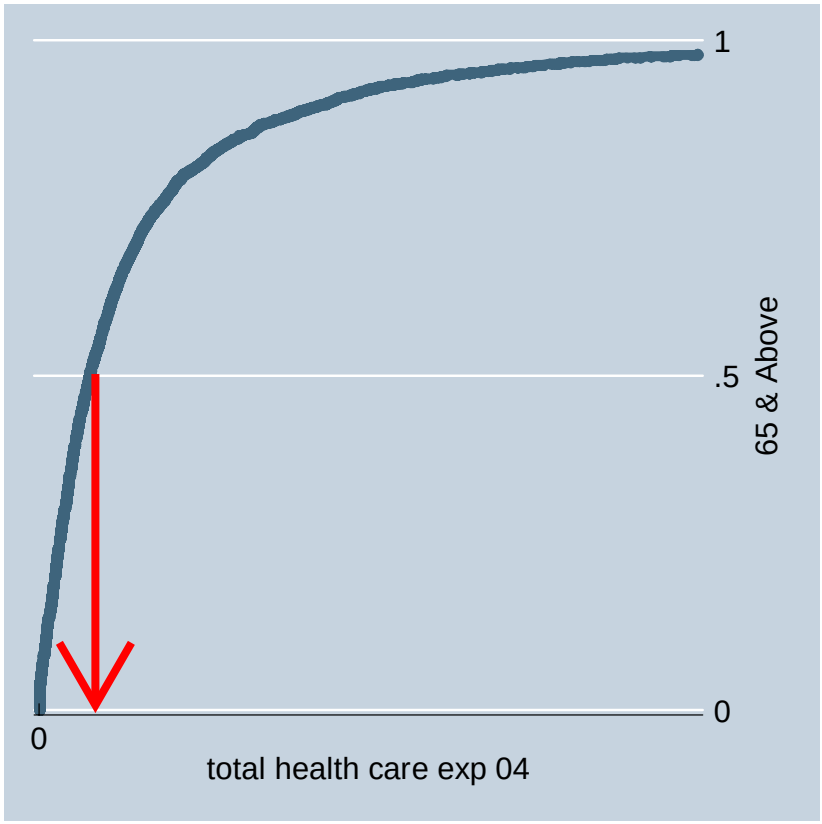




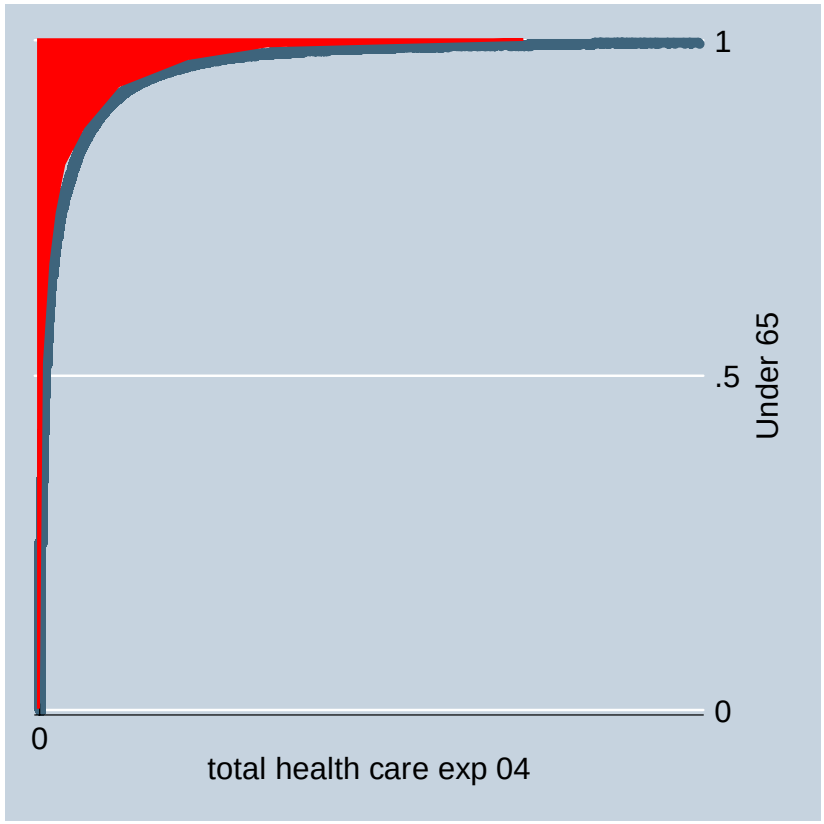


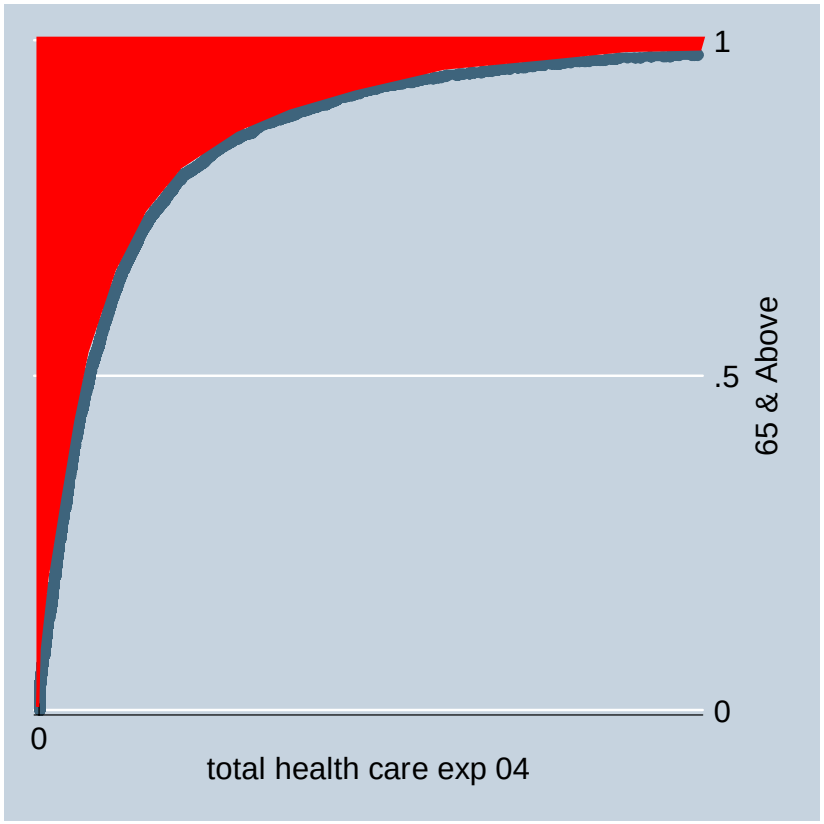
THE UNIVERSITY of WISCONSIN MADISON





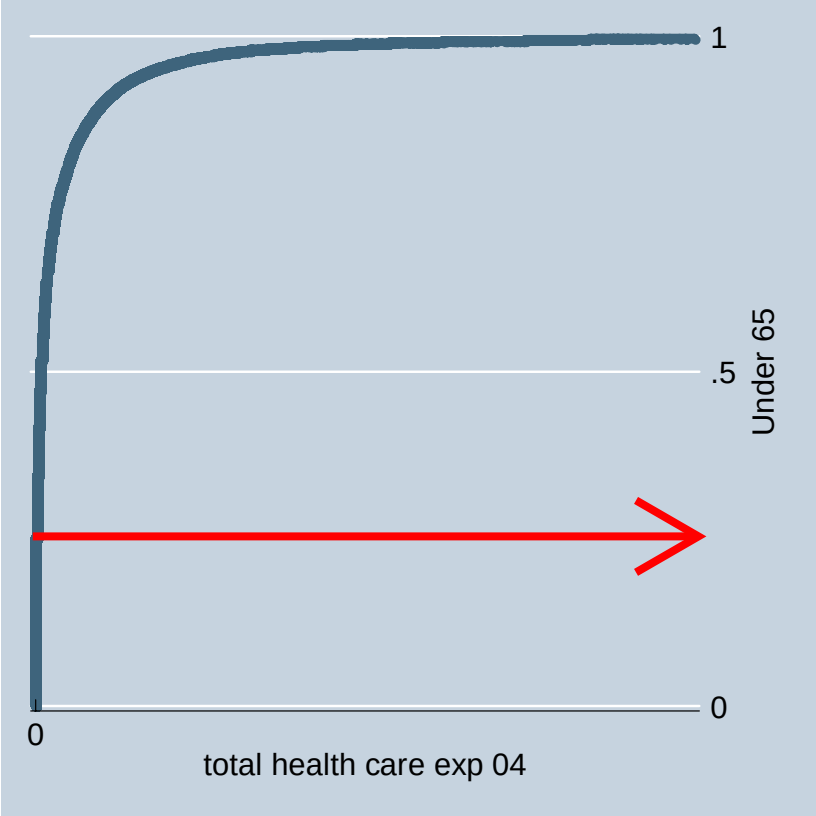
Conditional Quantiles

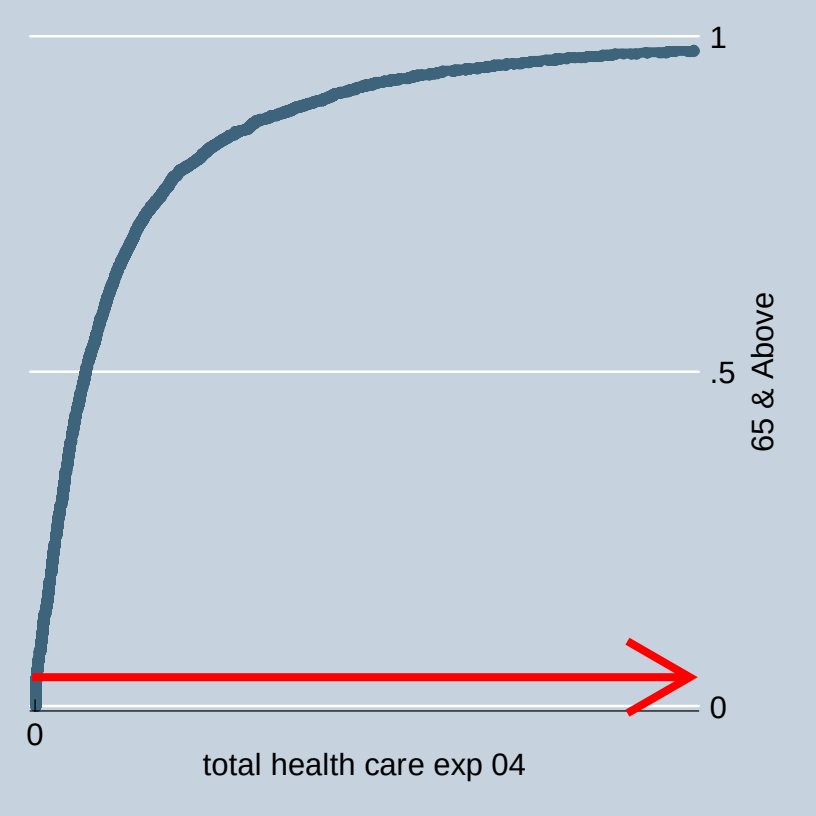




Conditional Means

THE UNIVERSITY of WISCONSIN MADISON





Conditional Probabilities

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

What are the Data, What are the Questions, and How Can Data Address the Questions?

* Data are given by sampling from the "anatomy" of the cost & covariate distribution $\phi(y, x)$

* Prominent data considerations and features:

  - $y \geq 0$ with sometimes nontrivial # $y = 0$

  - Conditional on x or not, distributions of observed y are often skewed

  - Timeframe: All relevant y observed, or possible left- &/or right-censoring?

  - Parametric analysis sets $\phi(y, x) = \phi(y, x; \theta)$

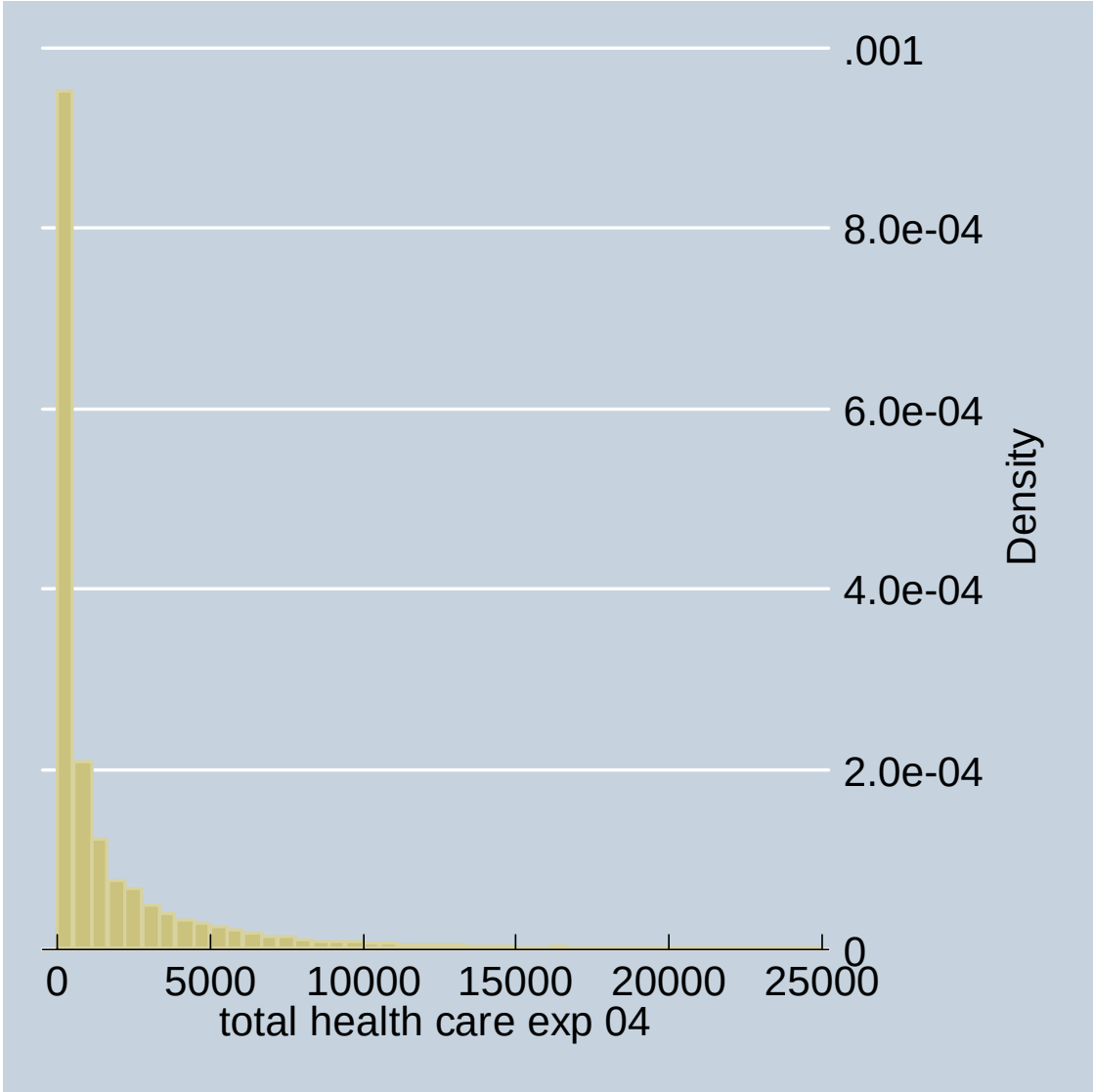
Questions of concern to decision makers, when translated into statistical terms, are typically based on properties of $\phi(y)$ or $\phi(y \mid x)$:

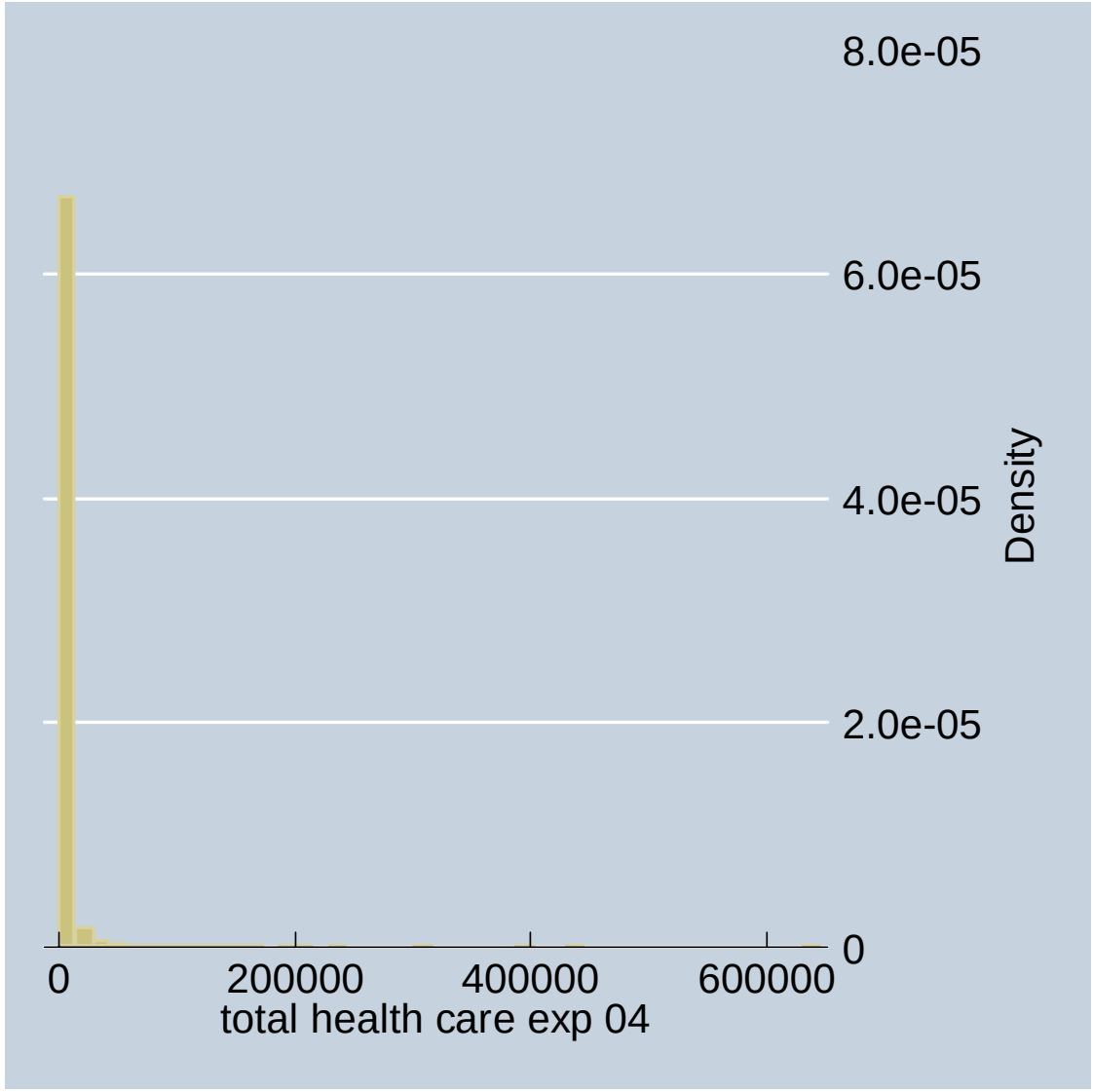
- $E[y] = \mu(= \mu(\theta))$ **or** $E[y \mid x] = \mu(x)$
- $Prob(y \in S) = \pi$ **or** $Prob(y \in S \mid x) = \pi(x)$
- **"Partial effects" of the x's on the above conditionals**
- **etc.**

*** Scientifically logical (and maybe regulatorily necessary) that specification of the parameter(s) of interest is *prior* to the analysis**

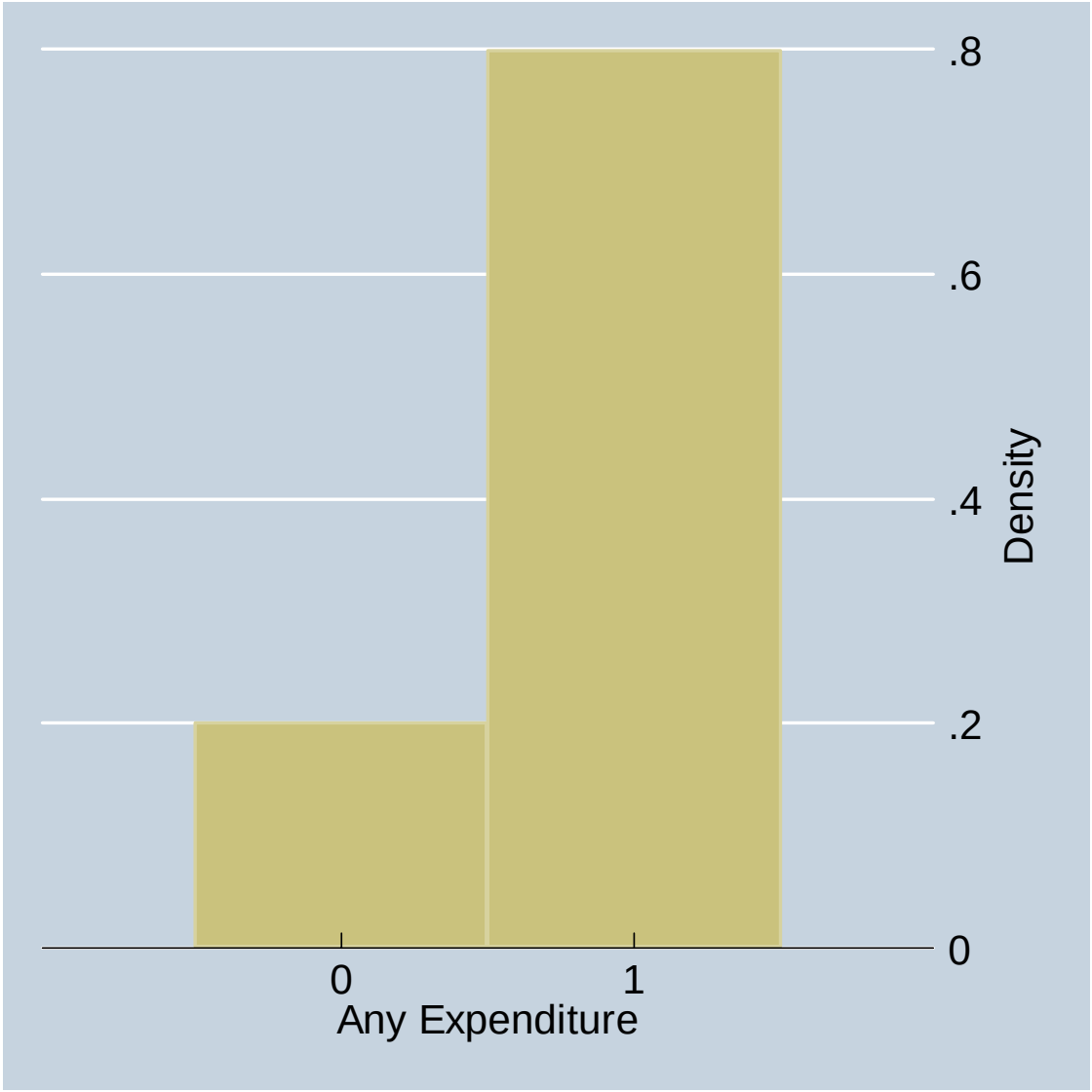
*** Decisionmaker's "loss function" will weigh considerations of bias, precision, etc. *given* criterion parameter(s) to be estimated**

THE UNIVERSITY of WISCONSIN MADISON





THE UNIVERSITY of WISCONSIN MADISON



THE UNIVERSITY of WISCONSIN MADISON

MEPS Total Health Care Expenditures 2004

	Percentiles	Smallest		
1%	0	0		
5%	0	0		
10%	0	0	Obs	34403
25%	68	0	Sum of Wgt.	34403
50%	500		Mean	2871.412
		Largest	Std. Dev.	9417.044
75%	2342	314498		
90%	6789	392462	Variance	8.87e+07
95%	12266	440524	Skewness	20.90473
99%	37560	645980	Kurtosis	962.3038

MEPS Total Health Care Charges 2004 (excl. Rx)

	Percentiles	Smallest		
1%	0	0		
5%	0	0		
10%	0	0	Obs	34403
25%	67	0	Sum of Wgt.	34403
50%	472		Mean	4887.386
		Largest	Std. Dev.	21636.52
75%	2274	742110		
90%	9825	859108	Variance	4.68e+08
95%	21495	1005849	Skewness	20.71001
99%	76815	1305474	Kurtosis	791.3738

THE UNIVERSITY of WISCONSIN MADISON

Modeling (Conditional) Means

* Probably the main focus in applied econometric work in this area is on modeling of conditional means $E[y \mid x] = \mu(x)$

  * Even though focus may sometimes be on marginal means E[y], estimates of these are recoverable from conditional mean estimates in conjunction with x-weighting

  * Given $y \geq 0$, then (*if it exists*) the conditional mean will satisfy $\mu(x) > 0$ except in trivial cases

Central (and Perhaps Decision-Relevant) Considerations in Conditional Mean Estimation

* Knowing $E[y \mid x]$ vs. knowing "factors" of $E[y \mid x]$, e.g.

$$E[y \mid x] = \Pr(y > 0 \mid x) \times E[y \mid y > 0, x]$$

  - "2PM" vs. "1PM", e.g.

* Enforcing vs. not enforcing constraints such as $\mu(x) > 0$

  - Linear- vs. log-link functions, e.g.

* Focus on $E[y \mid x]$ vs. focus on partial x-effects of $E[y \mid x]$

* Emphasis on consistent estimation vs. emphasis on estimator precision

  - Transform / retransform vs. levels estimation
  - Link function focus vs. family / distribution focus

THE UNIVERSITY of WISCONSIN MADISON

(cont.)

* **Focus on "main effects" vs. focus on "interaction effects"**
    - **Policy considerations (e.g. comorbidities)**
    - **Interpretational considerations (What is an IE?)**
    - **Sample size considerations and "overfitting"**
* **Endogenous vs. exogenous covariates**
    - **Rationales for including vs. excluding endogenous covariates (e.g. risk adjustment exercises)**
    - **Implications of possible endogeneity bias vis-a-vis the decision problem at hand**
    - **Strategies for mitigating endogeneity bias**

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

Goodness of Fit Considerations and Tests: What Departures of Model from Data are of Concern?

Means (conditional)

* Conditional moment-type tests

  - Focus on particular departures of concern in x-space
  - Hosmer-Lemeshow-, RESET-, G-O-L-, Whang-type tests

Distributions (conditional) more generally

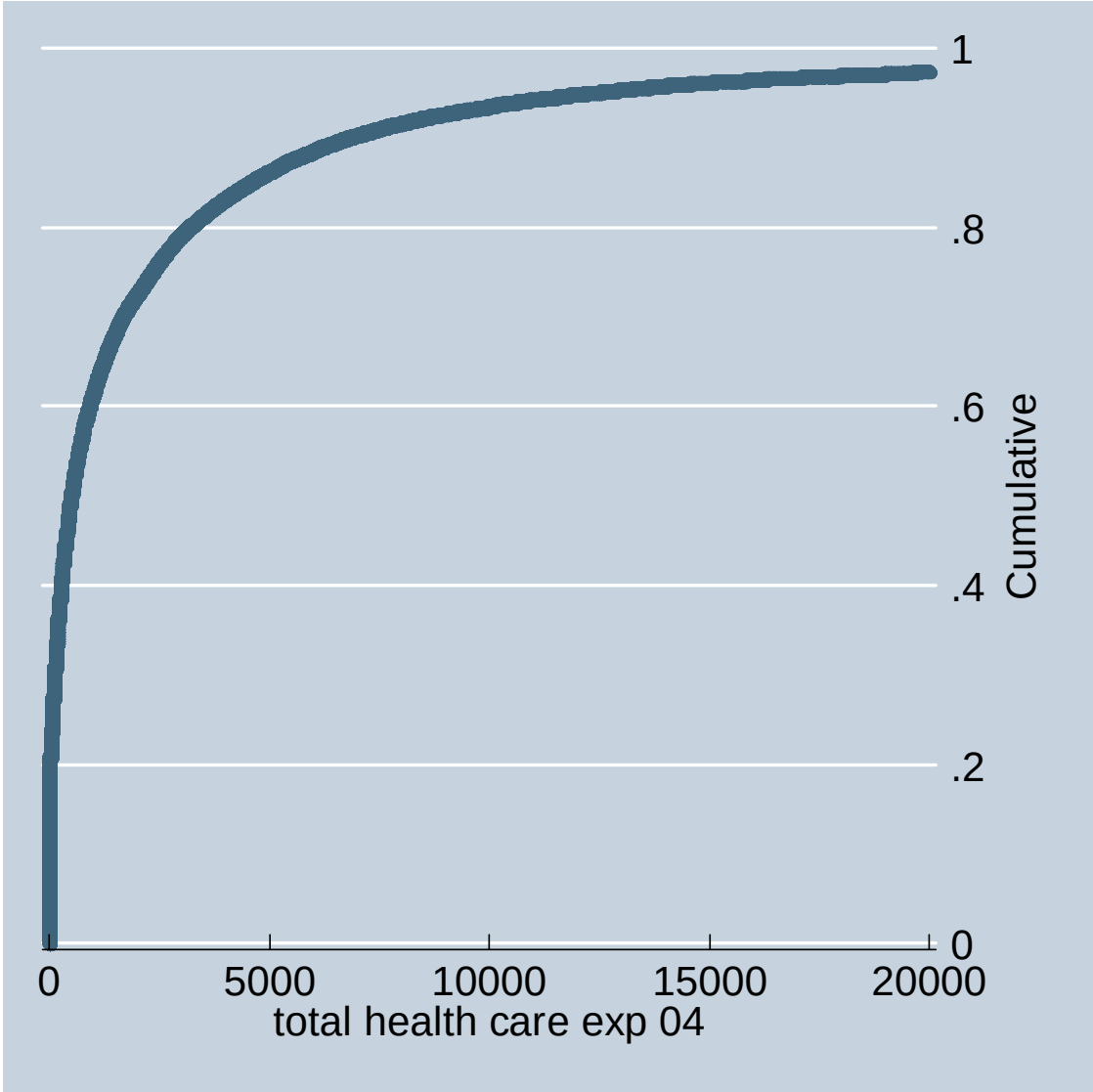
* Classical chi-square cell predictions

  - Focus on particular cells of concern

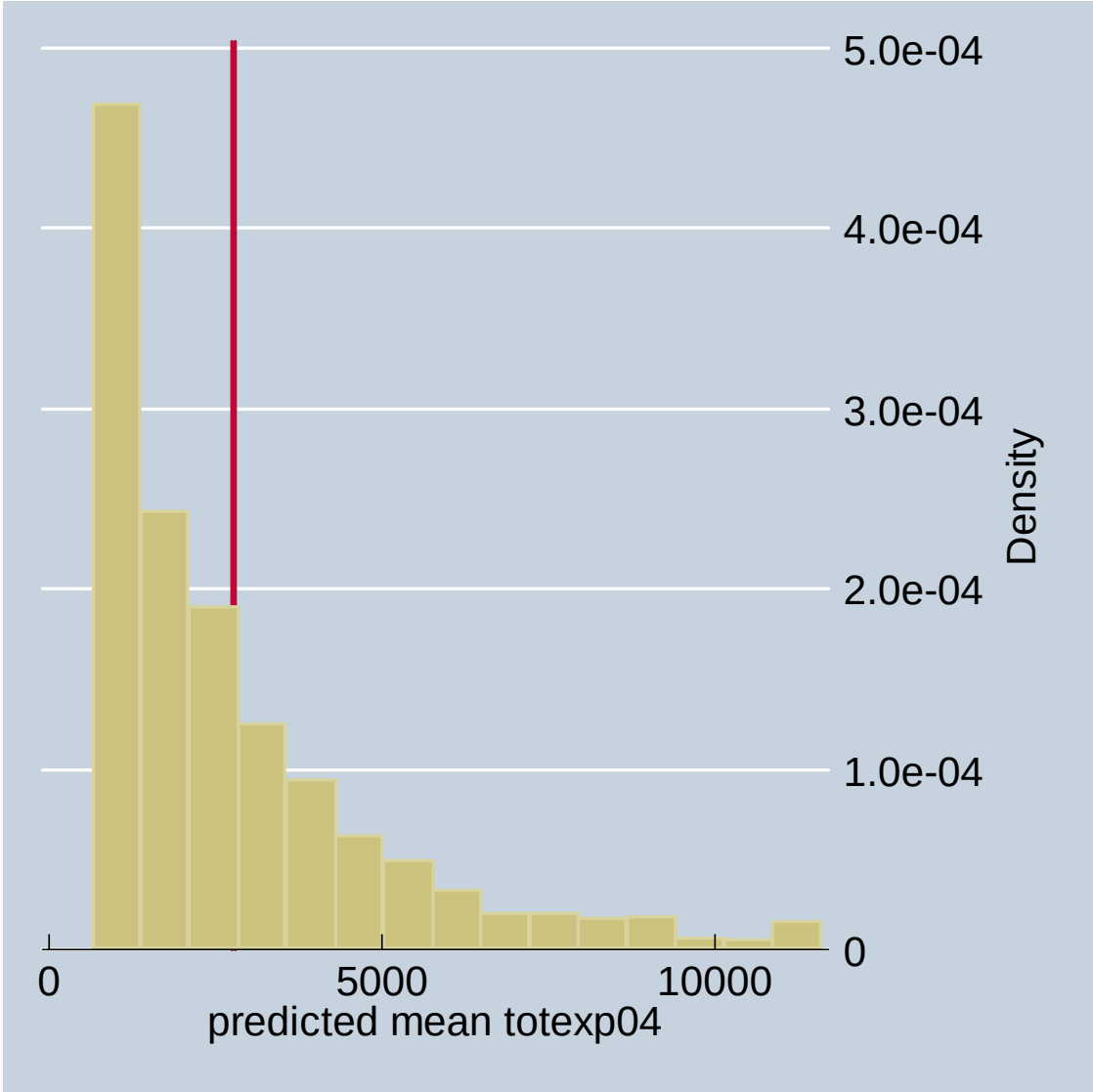
* Kolmogorov-Smirnov-type

In-Sample vs. Out-of-Sample "Fit" ("Risk Adjustment")

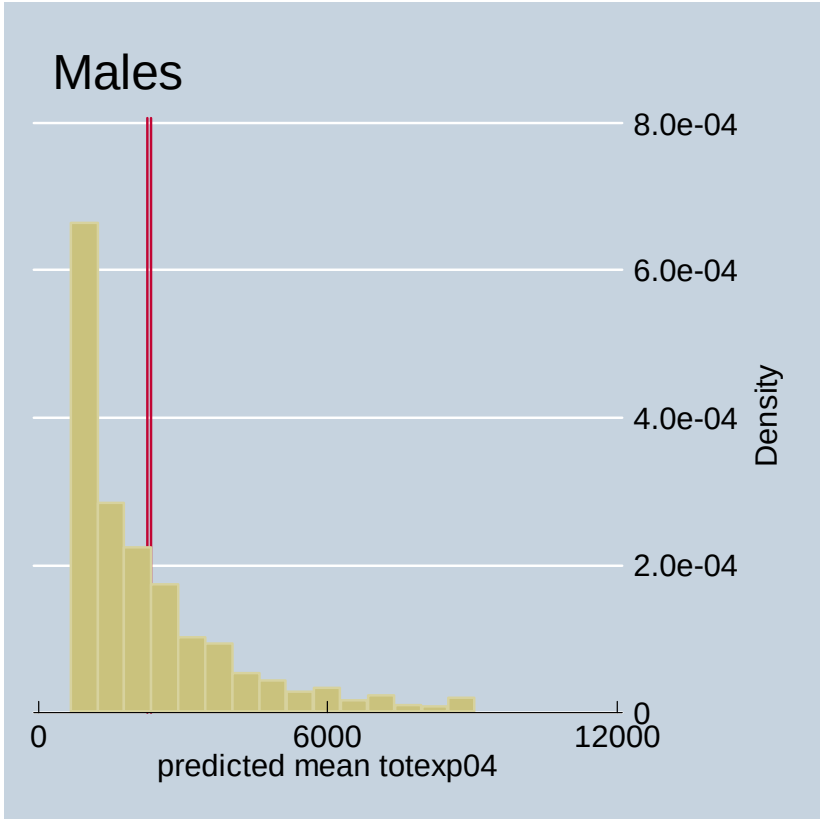
THE UNIVERSITY of WISCONSIN MADISON

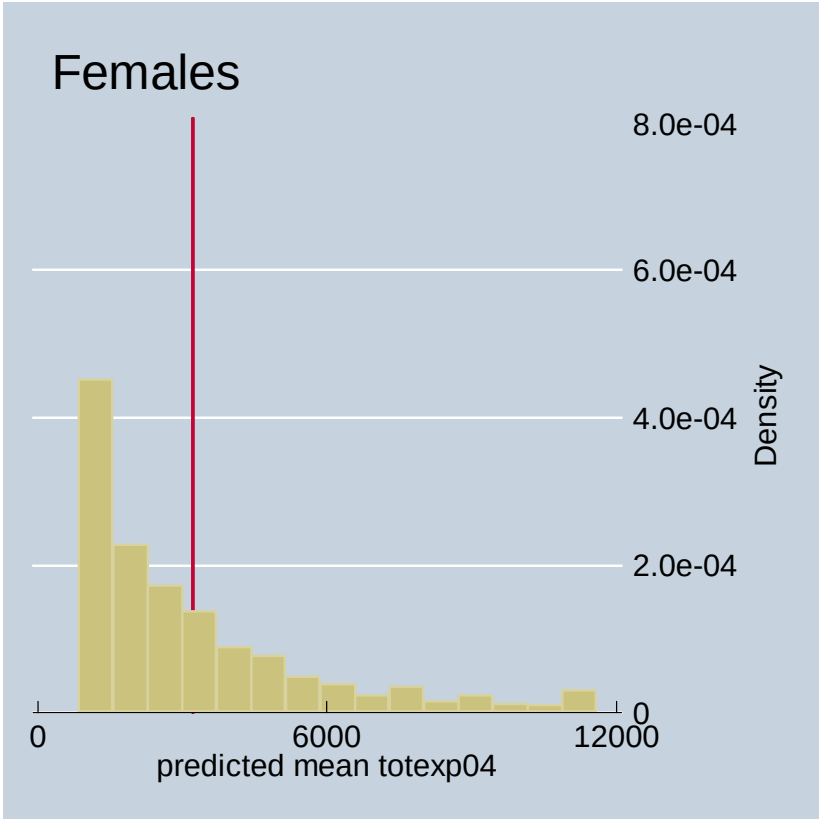


THE UNIVERSITY of WISCONSIN MADISON



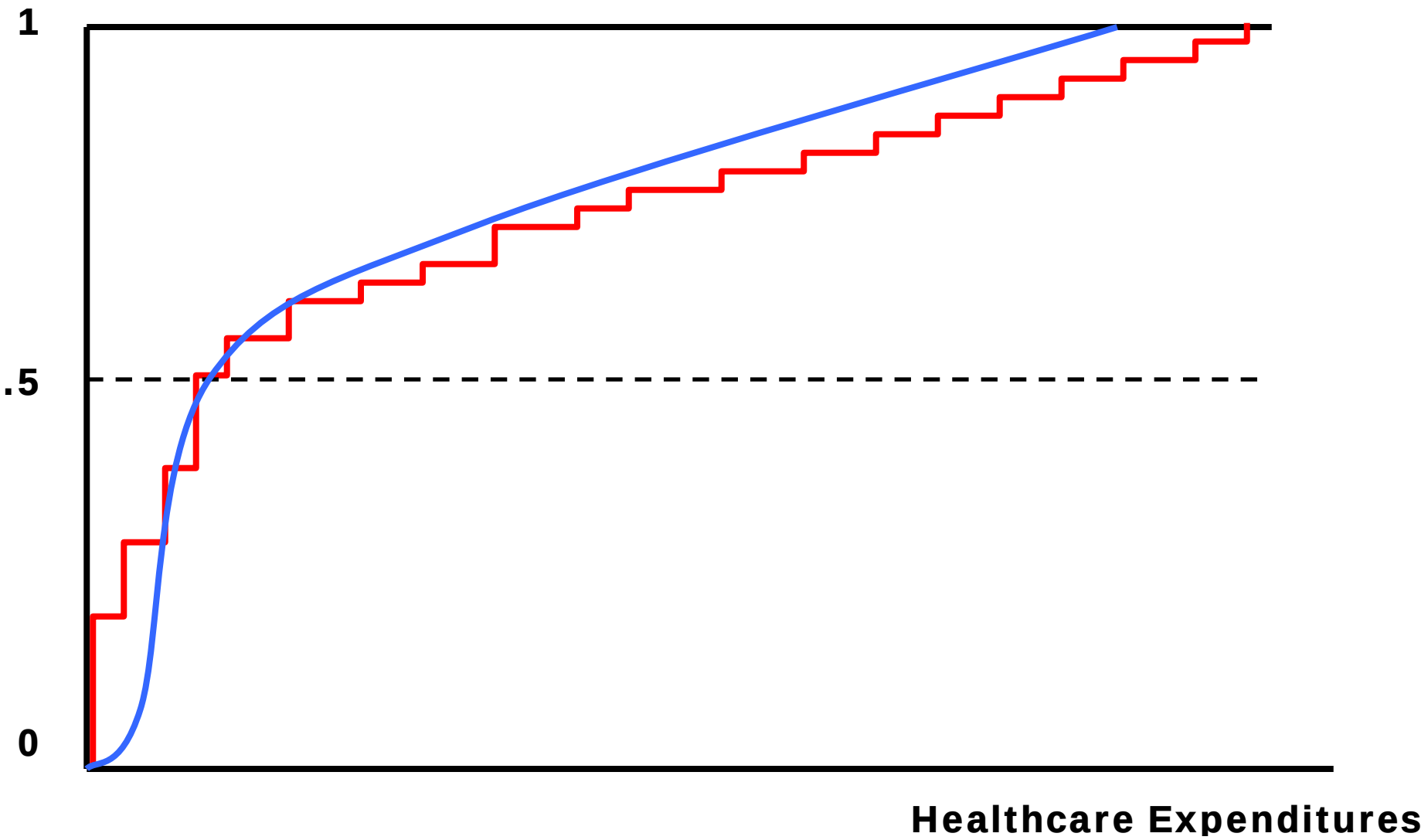
THE UNIVERSITY of WISCONSIN MADISON





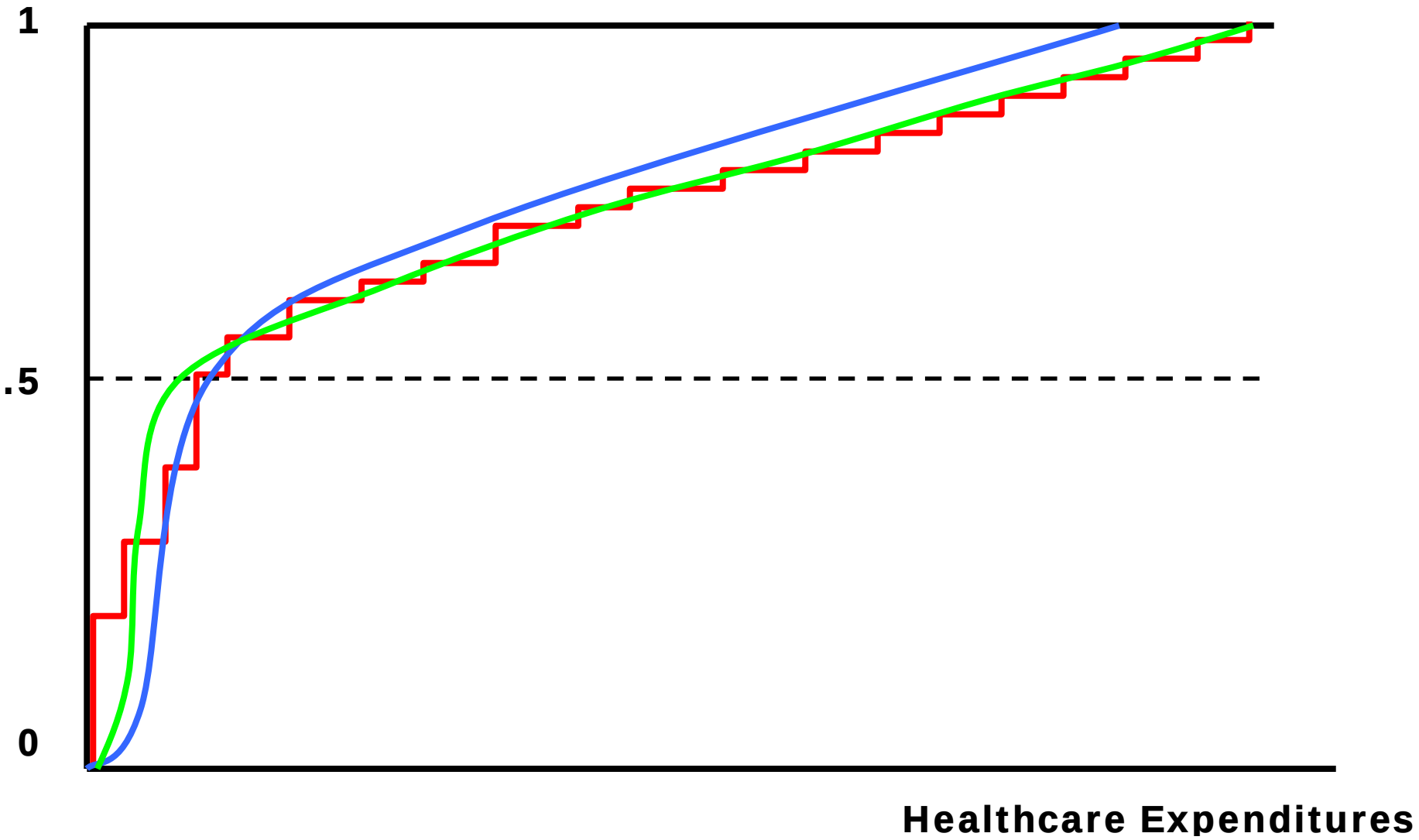
THE UNIVERSITY of WISCONSIN MADISON

Goodness of Fit Considerations: Mean? Median? Tails?



THE UNIVERSITY of WISCONSIN MADISON

Goodness of Fit Considerations: Mean? Median? Tails?



THE UNIVERSITY of WISCONSIN MADISON

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

Upper Tail Behavior

* The often extreme skewness of marginal and/or conditional cost distributions is frequently treated as an analytical nuisance

  * In some instances, high-end ("influential") cases are effectively downweighted via strategies like robust regression, quantile regression, or "outlier" elimination (trimming)

  * In such cases, however, analysts must obviously be attuned to what parameters are being estimated after such methodologies are applied, *and* how these relate to the decision problem at hand (e.g. means vs. medians)

ARCHIVES OF
PEDIATRICS
& ADOLESCENT MEDICINE

Vol. 156 No. 5, May 2002
Article

TABLE OF CONTENTS >

Health Care Use and Costs for Children With Attention-Deficit/Hyperactivity Disorder

National Estimates From the Medical Expenditure Panel Survey

Eugenia Chan, MD, MPH; Chunliu Zhan, MD, PhD;
Charles J. Homer, MD, MPH

Arch Pediatr Adolesc Med. 2002;156:504-511.

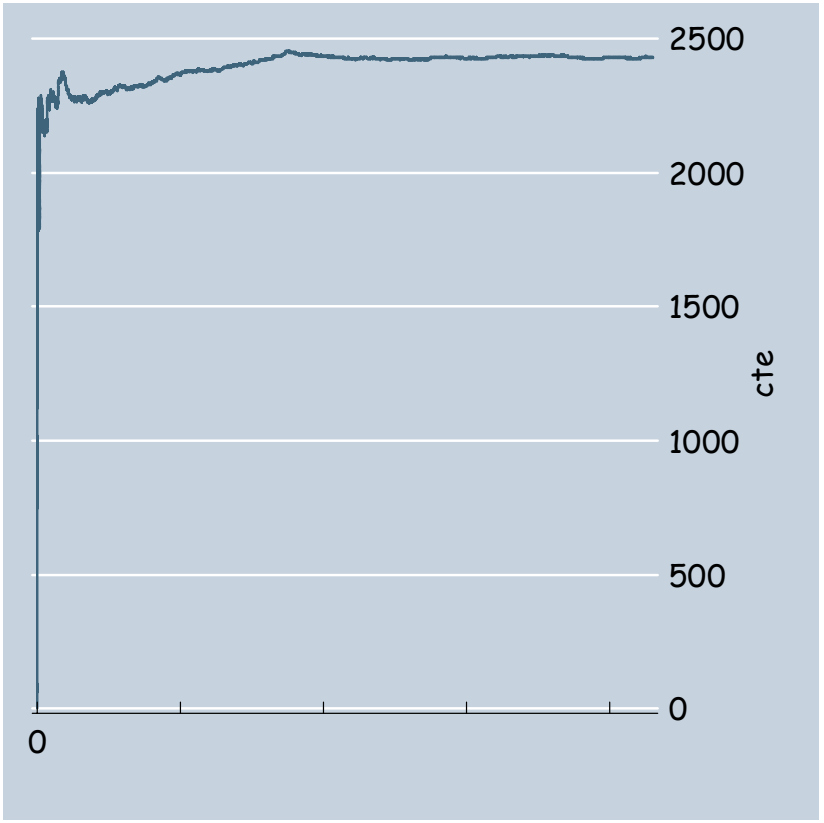
... Although expenditure data are typically somewhat skewed, the weighting involved in national survey data allows for the application of the central limit theorem and the use of means-based tests instead of either nonparametric tests or log transformation of the data. We were thus able to preserve the interpretability of the results as dollars. We examined expenditure distributions to identify extreme outliers that could drive the difference in means and excluded 3 children whose total expenditures exceeded $1 million. Because these 3 children all belonged in the general population group, excluding them as outliers slightly increased the estimated excess use and costs for the ADHD and asthma groups.

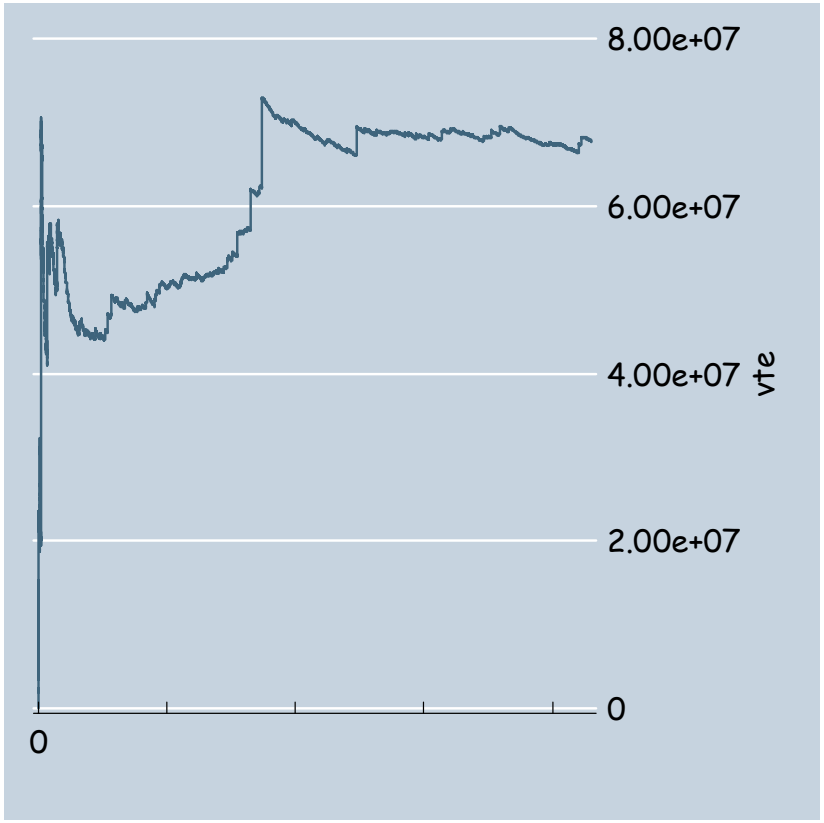
THE UNIVERSITY of WISCONSIN MADISON

Beyond such considerations, however, considerations of upper-tail behavior may be of interest:

*** From policy perspectives -- e.g. insurance / reinsurance, disease management, provider profiling, etc. -- the behavior of the upper α – percentile of the (conditional) cost distribution (and, consequently, how to model it) may be a primary consideration**

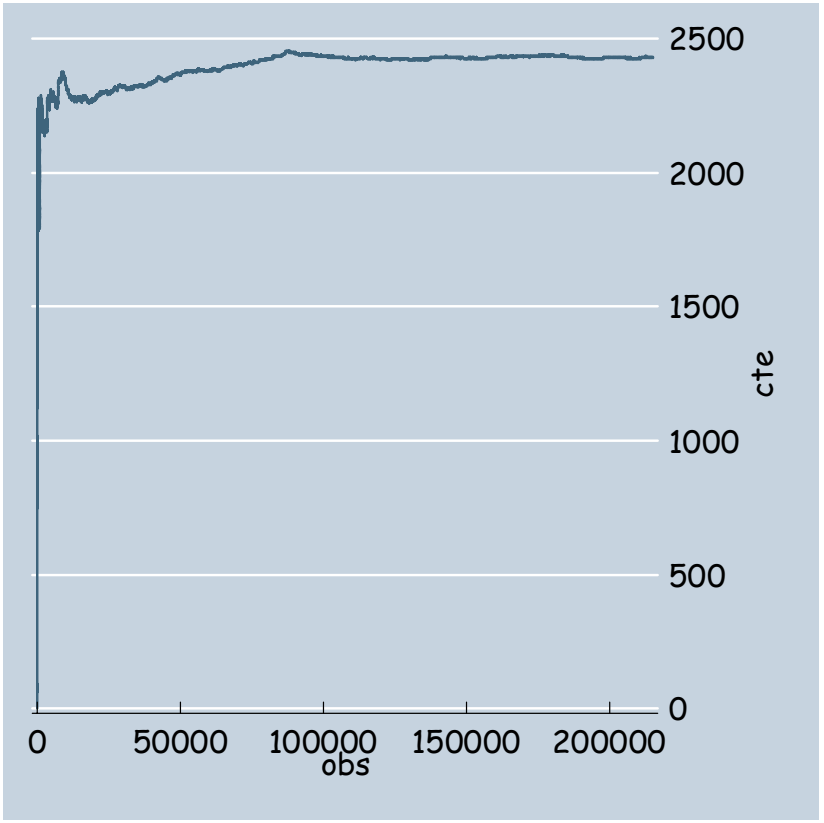
*** From a perspective of estimation and inference, the phenomena of heavy upper tails raise prospects of "Pareto-type" statistical behaviors that may require special attention**

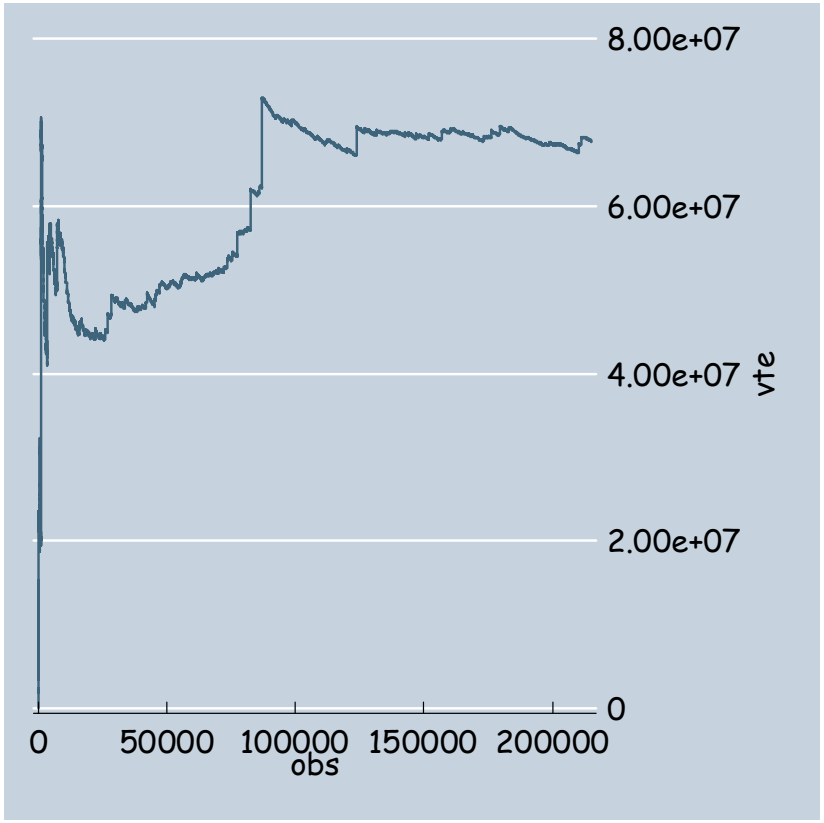




MEPS Total Expenditures: Running Mean & Variance

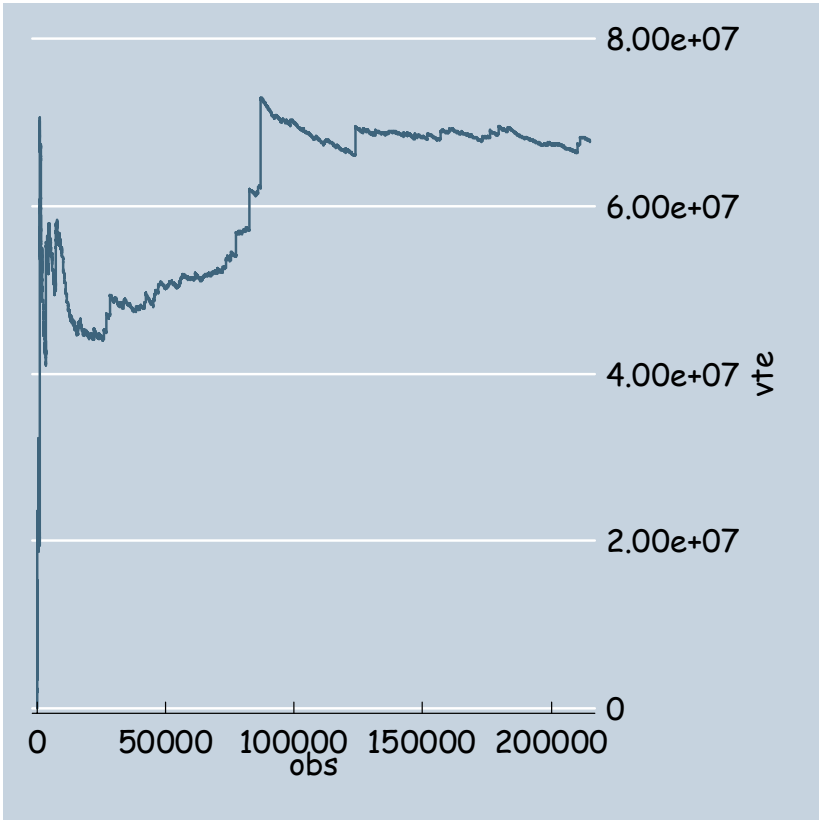
THE UNIVERSITY of WISCONSIN MADISON

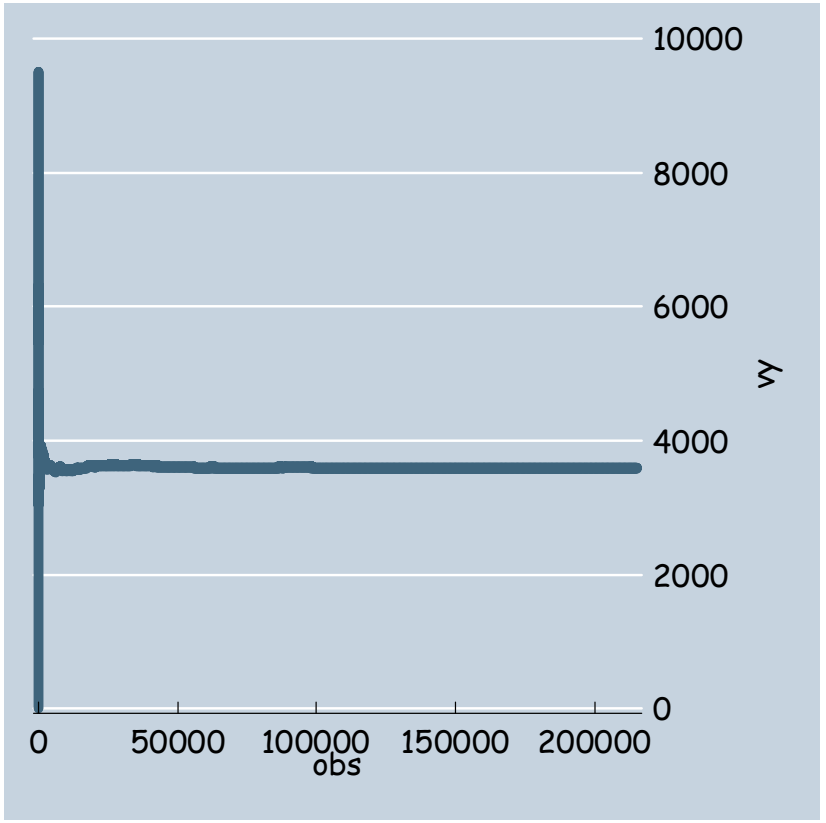




MEPS 1998-2004, N = 215125

THE UNIVERSITY of WISCONSIN MADISON





MEPS vs. Gaussian Running Variance

THE UNIVERSITY of WISCONSIN MADISON

THE JOURNAL OF
POLITICAL ECONOMY

Volume LXXI OCTOBER 1963 *Number 5*

NEW METHODS IN STATISTICAL ECONOMICS[1]

BENOIT MANDELBROT
Harvard University and I.B.M. Corporation

ROADMAP

1. How is empirical modeling useful in practice?

2. What sorts of questions may usefully be informed by such empirical investigations?

3. Specification and estimation of statistical models, in light of (2) and of the data

4. Goodness-of-Fit: Why and how?

5. Considerations of "high-end" costs

6. Translating and reporting research usefully

Standardization and Reporting

* Possible limits of standardization given heterogeneity of decision problems or questions confronting *users* of econometric studies

* Potential benefits of a standardized taxonomy to which producers and users of such research can point when defining terms, methods, etc.

* **BUT** emphasize that...

  ... work that doesn't fit (conceptually, methodologically, etc.) within the taxonomy can nonetheless be of high scientific merit

  ... work that does fit within the taxonomy is not necessarily of high scientific merit

* Well-conceived graphical presentation of data and results offers much promise in the knowledge transfer domain

THE UNIVERSITY of WISCONSIN MADISON