

**Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Government Agency Application for Data**

This application is to be used by all applicants, except Government Agencies, as defined in 957 CMR 5.02.

NOTE: In order for your application to be processed, you must submit the required application fee. Please consult the fee schedules for APCD and Case Mix data for the appropriate fee amount. A remittance form with instructions for submitting the application fee is available on the CHIA [website](#).

I. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name:	Keith Marzilli Ericson (primary applicant)
Title:	Assistant Professor of Markets, Public Policy, & Law and Faculty Research Fellow
Organization:	Boston University School of Management and National Bureau of Economic Research
Co-Investigator:	Jim Rebitzer
Title:	Professor of Management, Economics and Public Policy; Everett J. Lord Distinguished Scholar; Research Associate
Organization:	Boston University School of Management and National Bureau of Economic Research
Co-Investigator:	Benjamin Lubin
Title:	Assistant Professor of Information Systems
Organization:	Boston University School of Management
Co-Investigator:	Brigham Frandsen
Title:	Assistant Professor of Economics
Organization:	Brigham Young University
Co-Investigator:	Kimberley Geissler
Title:	Research Associate
Organization:	Boston University School of Management
Co-Investigator:	Amanda Starc
Title:	Assistant Professor of Health Care Management
Organization:	University of Pennsylvania
Project Title:	Understanding Insurance, Provider Networks, and Outcomes
Date of Application:	October 2014
Project Objectives (240 character limit)	We examine characteristics of provider networks and insurance policies and relationships with patient outcomes to better understand provider and enrollee behaviors. We use regression techniques, network analysis, and simulation.
Project Research Questions	We investigate how consumers value insurance plan designs and provider

	<p>networks, and examine the links between plan design, network structure, and outcomes. Specifically, we ask the following questions:</p> <ol style="list-style-type: none"> 1. What are the consequences of broader versus narrower insurance plan choice set (i.e. variation in deductibles, actuarial value, etc.)? 2. Can more complex insurance contracts improve outcomes and patient welfare by linking cost-sharing to more information (e.g. provider quality, patient health status)? 3. What is consumer willingness-to-pay for additional network access from their employer’s plan menu, and how does this affect insurance plan design? 4. Do consumers with greater medical utilization gravitate towards certain kinds of plans or networks of providers? 5. What do professional networks of shared patients among physicians look like, and how do such networks vary by type of insurance plan (e.g. HMO v PPO vs. Medicaid)? 6. What is the relationship between professional networks of shared patients among physicians, resource use, and patient outcomes?
--	--

I. PROJECT SUMMARY

Briefly describe the purpose of your project and how you will use the requested CHIA data to accomplish your purpose.

This project investigates how consumers value insurance plans and networks. This requires we establish links between plan design, network structure, and health/utilization outcomes. We will use the CHIA data to first examine the link between characteristics of insurance plans and the utilization of health care, including particular procedures and total spending. We then examine the provider networks available to consumers in each plan/insurance type. We estimate models of demand for particular providers (e.g. hospitals, using measures of geographical distance) and then for insurer-specific networks. We also examine how network structure is associated with outcomes. We associate network-breadth measures with price levels, utilization patterns, and enrollee composition. We also model the structure of provider-provider connections (e.g. referral networks) using insights from network theory (e.g. concepts of connectedness, centrality, etc). We examine how this structure varies by insurance plan design, and how these structures are associated with measures of health outcomes and process quality.

These analyses will use standard forms of regression analysis, hazard models, simulated method of moments, models of consumer choice (e.g. differentiated product demand models) and welfare (e.g. expected utility models), and network structure modeling (e.g., clustering coefficients, betweenness, and spectral analysis).

II. FILES REQUESTED

Please indicate the databases from which you seek data, the Level(s) and Year(s) of data sought.

ALL PAYER CLAIMS DATABASE	Level 1 ¹ or 2 ²	Single or Multiple Use	Year(s) Of Data Requested Current Yrs. Available 2009 - 2012

¹ Level 1 Data: De-identified data containing information that does not identify an individual patient and with respect to which there is no reasonable basis to believe the data can be used to identify an individual patient. This data is de-identified using standards and methods required by HIPAA.

² Level 2 (and above) Data: Includes those data elements that pose a risk of re-identification of an individual patient.

<input checked="" type="checkbox"/> Medical Claims	<input type="checkbox"/> Level 1 ³ <input checked="" type="checkbox"/> Level 2	Select... e Singl	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012
<input checked="" type="checkbox"/> Pharmacy Claims	<input checked="" type="checkbox"/> Level 2	Select... e Singl	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012
<input type="checkbox"/> Dental Claims	<input type="checkbox"/> Level 2	Select... e Singl	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012
<input checked="" type="checkbox"/> Member Eligibility	<input checked="" type="checkbox"/> Level 2	Select... e Singl	
<input checked="" type="checkbox"/> Provider	<input checked="" type="checkbox"/> Level 2	Select... e Singl	
<input checked="" type="checkbox"/> Product	<input checked="" type="checkbox"/> Level 2	Select... e Singl	

CASEMIX	Level 1 - 6	Fiscal Years Requested
Inpatient Discharge	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures <input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	<u>1998-2013 Available</u> (limited data 1989-1997)
Outpatient Observation	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures <input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	<u>2002-2012 Available</u> (2013 available 8/1/14)
Emergency Department	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN; Stated Reason for Visit <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures	<u>2000-2012 Available</u> (2013 available 9/1/14)

³ Please note that Level 1 APCD data is not available as of 4/30/2014. This is scheduled to be available later in 2014.

	<input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	
--	---	--

III. FEE INFORMATION

Please consult the fee schedules for APCD ([Administrative Bulletin 13-11](#)) and Case Mix data ([Administrative Bulletin 13-09](#)) and select from the following options:

APCD Applicants Only

- Academic Researcher
- Others (Single Use)
- Others (Multiple Use)

Case Mix Applicants Only

- Single Use
- Limited Multiple Use
- Multiple Use

Are you requesting a fee waiver?

- XXX Yes
- No

If yes, please submit a letter stating the basis for your request. Please refer to the fee schedule for qualifications for receiving a fee waiver. If you are requesting a waiver based on the financial hardship provision, please provide documentation of your financial situation. Please note that non-profit status alone isn't sufficient to qualify for a fee waiver.

IV. REQUESTED DATA ELEMENTS [APCD Only]

State and federal privacy laws limit the use of individually identifiable data to the minimum amount of data needed to accomplish a specific project objective. Please use the [APCD Data Specification Workbook](#) to identify which data elements you would like to request and attach this document to your application.

V. MEDICAID DATA [APCD Only]

Please indicate here whether you are seeking Medicaid Data:

- Yes
- No

Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are directly connected with the administration of the Medicaid program. If you are requesting Medicaid data from Level 2 or above, please describe in detail why your use of the data meets this requirement. Applications requesting Medicaid data will be forwarded to MassHealth for a determination as to whether the proposed use of the data is directly connected to the administration of the Medicaid program. MassHealth may impose additional requirements on applicants for Medicaid data as necessary to ensure compliance with federal laws and regulations regarding Medicaid.

This project will identify how insurance plans can be designed to improve individuals' health outcomes
--

and welfare, and how to increase health system efficiency. The results of our analyses will inform market regulators as they evaluate policies, particularly those that affect network coverage in MassHealth. Our results will identify patterns of provider networks that patients value and that deliver effective care, giving more information to the Medicaid program about efficient allocation of providers. Additionally, many studies have examined the effects of limited provider networks for Medicaid – we will look at the structure of these networks and their relationships with patient outcomes, particularly as compared to other types of insurance including HMOs. This will potentially inform Medicaid as to the value of expanding networks (if any) in terms of outcomes including cost and utilization measures such as hospitalizations or emergency department visits.

VI. REQUESTS PURSUANT TO 957 CMR 5.04

If you are a payer, provider, provider organization or researcher seeking access to Level 1 (de-identified) data, please describe how you will use such data for the purposes of lowering total medical expenses, coordinating care, benchmarking, quality analysis or other administrative research purposes. Please provide this information below.

VII. FILTERS

If you are requesting APCD elements from Level 2 or above, describe any filters you are requesting to use in order to limit your request to the minimum set of records necessary to complete your project. (For example, you may only need individuals whose age is less than 21, claims for hospital services only, or only claims from small group projects.)

APCD FILE	DATA ELEMENT(S) FOR WHICH FILTERS ARE REQUESTED	RANGE OF VALUES REQUESTED
Medical Claims		
Pharmacy Claims		
Dental Claims		
Membership Eligibility		
Provider		
Product		

VIII. PURPOSE AND INTENDED USE

1. Please explain why completing your project is in the public interest.

This project will identify how insurance plans can be designed to improve individuals' health outcomes and welfare, and how to increase health system efficiency. The results of our analyses will inform market regulators as they evaluate policies that affect network coverage and plan generosity in the health insurance exchanges, as well as in the individual and group markets. Our results will identify patterns of provider networks that patients value and that deliver effective care, which will be of use to providers and insurance plans/carriers that desire to improve care.

2. **Attach** a brief (1-2 pages) description of your research methodology. (This description will not be posted on the internet.)
3. Has your project received approval from your organization’s Institutional Review Board (IRB)? Please note that CHIA will not review your application until IRB documentation has been received (if applicable).
 - X Yes, and a copy of the approval letter is attached to this application.
 - No, the IRB will review the project on _____.
 - No, this project is not subject to IRB review.
 - No, my organization does not have an IRB.

IX. APPLICANT QUALIFICATIONS

1. Describe your qualifications to perform the research described or accomplish the intended use of CHIA data.

Keith Marzilli Ericson holds a PhD in Economics from Harvard University and a BA in economics and political science from Williams College. He is an Assistant Professor of Markets, Public Policy, and Law in the Boston University School of Management, teaching courses in econometrics and optimization theory. He is also a Faculty Research Fellow at the National Bureau of Economic Research.

Jim Rebitzer holds a Ph.D. in economics from the University of Massachusetts at Amherst and a BS in biology from the University of Illinois in Urbana Illinois. He is a professor of Management, Economics and Public Policy at the Boston University School of Management where he is also an Everett J. Lord distinguished scholar as well as being a research associate at the National Bureau of Economic Research. He has considerable experience working with confidential commercial insurance records in empirical, health services research.

Benjamin Lubin holds a PhD in Computer Science from Harvard University and an AB in Computer Science from Harvard University. He is an Assistant Professor of Information Systems in the Boston University School of Management, teaching courses in computer science. He is also a Hariri Institute junior faculty fellow.

Brigham Frandsen holds a PhD in Economics from the Massachusetts Institute of Technology and a BS in economics and physics from Brigham Young University. He is an Assistant Professor of Economics at Brigham Young University, teaching courses in econometrics. He also was a Robert Wood Johnson Scholar in Health Policy Research at Harvard University.

Kimberley Geissler holds a PhD in Health Policy and Management from the University of North Carolina Gillings School of Global Public Health and a BA in chemistry and economics from Williams College. She is a Research Associate in the Boston University School of Management. She is also an Adjunct Assistant Professor of Health Policy and Management at the University of North Carolina.

Amanda Starc holds a PhD in Business Economics from Harvard University and a BA in Economics from Case Western Reserve. She is an Assistant Professor of Health Care Management in the Wharton School, University of Pennsylvania.

Ericson and Starc have examined the Massachusetts Health Insurance Exchange (Connector) in a series of papers (2013 American Economic Review, 2013 Inquiry, 2012 National Bureau of Economic Research

Working Paper, 2013 Working Paper). All investigators have worked with sensitive health data previously (Ericson: Marketscan; Rebitzer: numerous datasets including commercial insurance claims data; Frandsen: confidential medical and drug claims data from a large private insurer; MEPS Insurance Component; Census Long Form (confidential 1:6 sample); Lubin: Marketscan; Starc: Medigap; Geissler: State Hospital Discharge Datasets, Marketscan).

2. Attach résumés or curricula vitae of the applicant/principal investigator, key contributors, and of all individuals who will have access to the data. (These attachments will not be posted on the internet.)

X. DATA LINKAGE AND FURTHER DATA ABSTRACTION

1. Does your project require linking the CHIA Data to another dataset?

- Yes
 No

2. If yes, will the CHIA Data be linked to other patient level data or with aggregate data (e.g. Census data)?

- Patient Level Data
 Aggregate Data

3. If yes, please identify all linkages proposed and explain the reasons(s) that the linkage is necessary to accomplish the purpose of the project. Please be specific in describing vvhich data elements will be linked to outside datasets and how this will be accomplished.

4. We propose to link CHIA data to the following aggregate datasets, described here:

- 1) Hospital linkages – We will link hospitals to the American Hospital Association Annual Survey Database (AHA) for hospital characteristics; to the Medicare Hospital Compare dataset for quality and aggregate health outcome data. We will use information on the service provider name and location to identify hospitals from the medical claims data (MC027, MC028, MC029, MC030, MC031, MC033, MC034, MC035), linked with location information from the provider file based on the National Service Provider ID (MC026 or MC024 if MC026 is missing). We will link using the hospital name and location we derive from the APCD fields to hospital name and location in the AHA and Hospital Compare datasets. This is needed to describe provider networks and model patient choice of hospital.
- 2) Provider linkages – We will link providers to the American Medical Association Physician Masterfile for provider specialty and demographic data (using the MC026, National Service Provider ID; if MC026 is missing, we will use MC024 and the provider file to link to the AMA Masterfile by physician name and service location); to tiering (quality/cost-efficient care) measures for specialist providers participating in the GIC UniCare plans (using the MC026, MC028, MC029, MC030, and MC035 to link by National Provider ID if possible, or by name and location if not; if the information is not available in the medical

claims, we will use MC024 and the provider file to conduct the linkage); and to the Massachusetts Health Quality Partners (MHQP) provider dataset to accurately link providers to practices for determination of practice level measures and to link to quality data (ambulatory physician group practice linkages) [This linkage will be done using the MC026, National Service Provider ID; if MC026 is missing, we will use MC024 and the provider file to link to the MHQP provider dataset by physician name and service location]. To link to provider information, we need a number of provider identifiers. In published analyses and reports, we will not identify providers or report information where deductive disclosure would be possible (e.g., we will mask small cells, etc.). This is needed to characterize provider networks (e.g. are more efficient doctors more likely to be in the same network) and model patient choice of provider.

- 3) Ambulatory physician group practice linkages – We will link provider data to MHQP quality data on clinical and patient experience measures for primary care physicians (We will do this using provider IDs, names, and practice locations from the medical claims files and the provider file. We will use MC024, MC026, MC028, MC029, MC030, MC035 and the provider file to find the address and practice location if necessary for the linkage. We will use the physician group information from the MHQP provider dataset linkage). This is needed to characterize provider networks and model patient choice of provider.
- 4) Geographic area linkages – We will link member geographic data (ZIP, city, county from the member eligibility file – ME3, ME4, ME6, ME017) to the corresponding geographic indicator in the Area Resource File and the American Communities Survey/Census data to get information on healthcare supply, socioeconomic status, and regional characteristics. This is needed to account for variation in patient characteristics that might affect patient use of medical care or outcomes; we do not identify individual patients, merely link to characteristics of their ZIP code.
- 5) Carrier and/or insurance plan linkages – We will link carrier and/or insurance plan data (ME030, ME040 to link to product file, and then PR003, PR004) to market share and premium data from Mass Connector, as well as to the network definition of plans in the Connector (by looking up particular providers to determine if they are in-network for a given plan). This is needed to model insurer price setting (which is jointly determined with patient demand for insurance), how prices move with plan generosity, and to model consumer choice of insurance plan.

5. If yes, please identify the specific steps you will take to prevent the identification of individual patients in the linked dataset.

As these datasets do not increase the ability or likelihood of identification of individual patients in the linked datasets, these linkages would not jeopardize patient confidentiality. As discussed in the data security and integrity section, we will take great care to ensure the confidentiality of the data.

XI. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from such data, in any paper, report, website, statistical tabulation, seminar, conference, or other setting.

We will submit the results of our study for academic publication in peer-reviewed journals and present in seminars and conferences as necessary. We will have summary statistics and analyses completed using the data, but no identification of patients will be possible. If there are small cells in the analysis (<10 patients), we will censor these cells to maintain confidentiality.

2. Will the results of your analysis be publicly available to any interested party? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee.

The results will be available for no fee at the researchers' websites or upon email request.

3. Will you use the data for consulting purposes?

Yes
 No

4. Will you be selling standard report products using the data?

Yes
 No

5. Will you be selling a software product using the data?

Yes
 No

6. If you have answered "yes" to questions 3, 4 or 5, please describe the types of products, services or studies.

XII. USE OF AGENTS AND/OR CONTRACTORS

Third-Party Vendors. Provide the following information for all agents and contractors who will work with the CHIA Data.

Company Name:	
Contact Person:	
Title:	