Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Governmental Application for Case Mix Data

This form 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 schedule for the appropriate fee amount. A remittance form with instructions for submitting the application fee is available on the CHIA website.

. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name:	Amanda Cook
Title:	Researcher
Organization:	Purdue University
Project Title:	Market Power and Rate Setting for Hospitals and Insurance companies in MA
Mailing Address:	403 W State St West Lafayette, IN 47906
Telephone Number:	765 494.7675
Email Address:	ACook1@purdue.edu
Names of Co-Investigators:	James Bland
Email Addresses of Co-Investigators:	Bland1@purdue.edu
Original Data Request Submission Date:	2/24/2015
Dates Data Request Revised:	
Project Objectives (240 character limit)	This project aims to identify the determinants of bargaining power between insurers and hospitals and analyze the implications for welfare of various alternative policies and market structures.
Project Research Questions (if applicable)	1. How does market power of providers and insurance companies affect negotiated rates? 2. What determines the bargaining power of hospitals and insurers? 3. How is the surplus split between hospitals and insurers? 4. What are this implications for social welfare?

II. PROJECT SUMMARY

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

We aim to quantify and identify the determinants of bargaining power between insurance companies and hospitals. In order to achieve this, we will estimate negotiated rates between hospitals and insurance companies, and use Nash bargaining and other concepts from game theory to estimate bargaining power and division of surplus. We will run multivariate regressions to obtain estimates of

negotiated rates using total charges provided in the inpatient data coupled with publically provided data about insurance company per member per month costs. We will use these estimated models to asses the welfare implications of various changes in policy and market structure that could take place in the Massechusetts healthcare system.

III. FILES REQUESTED

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

CASE MIX	Levels 1 – 6	Fiscal Years Requested
Inpatient Discharge	☐Level 1 – No Identifiable Data Elements ☐Level 2 – Unique Physician Number (UPN) ☐Level 3 – Unique Health Information Number (UHIN) ☒Level 4 – UHIN and UPN ☐Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐Level 6 – Date of Birth; Medical Record Number; Billing Number **PLEASE PROVIDE JUSTIFICATION BELOW FOR REQUESTING THE CHOSEN LEVEL: We would like level 2 data because we need detailed procedure codes. I In addition, level two contains more demographic variables that we can use as control variables.** We need to identify geographic markets for health care, both from hospitals' and patients' perspective. We will use patient's zip codes, in conjunction with hospital utilization and insurance information to completely describe the market. Appropriately identifying hospitals which are substitutes is fundamental to this research project, and so we need five digit zip codes. This rational holds for Inpatients, Outpatients, and ER. Please see 'Outpatients' for justification of UPN Our research requires an aggregate merge between the APCD and Case Mix data bases. The required level of Case Mix data contains dates of services, and so we request level four data to be used in the merge.	Year 2013
Outpatient Observation	☐Level 1 – No Identifiable Data Elements ☐Level 2 – Unique Physician Number (UPN) ☐Level 3 – Unique Health Information Number (UHIN) ☒Level 4 – UHIN and UPN ☐Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐Level 6 – Date of Birth; Medical Record Number; Billing Number **PLEASE PROVIDE JUSTIFICATION BELOW FOR REQUESTING THE CHOSEN LEVEL: We would like level 2 data because we need detailed diagnosis codes. In addition, level two contains more demographic variables that we can use as control variables.** Unique Physician Number would be an important variable for two reasons: to link patient-level	Year 2013

	records from the APCD and Case Mix data sets, and to control for physicians different patterns of treatment or service usage. By having UPN, we could control for this variation in service which is unrelated to the patient's illness.	
Emergency Department	☐Level 1 – No Identifiable Data Elements ☐Level 2 – Unique Physician Number (UPN) ☐Level 3 – Unique Health Information Number (UHIN) ☒Level 4 – UHIN and UPN ☐Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐Level 6 – Date of Birth; Medical Record Number; Billing Number **PLEASE PROVIDE JUSTIFICATION BELOW FOR REQUESTING THE CHOSEN LEVEL: We would like level 2 data because we need detailed diagnosis codes. In addition, level two contains more demographic variables that we can use as control variables.** UPN and zip codes needed for the reasons stated above.	Year 2013

IV. FEE INFORMATION

Please consult the fee schedules for Case Mix data, available at http://chiamass.gov/regulations/#957_5, and select from the following options:

☒ Single Use
☐ Limited Multiple Use
☐ Multiple Use

Are you requesting a fee waiver?

☒ 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.

Both my co-author (James Bland) and I (Amanda Cook) are Ph.D. students at Purdue University. This research is student directed. A faculty member (Jack Barron) is listed on the IRB application, which is an institution requirement for Purdue, but the work will be student work. While the University is non-profit, PhD students are not allocated funding for data/research. Additionally, our stipends of $18,000/ year are sufficient for a basic standard of living, but are not sufficiently lavish to allow us to fund data acquisition.

V. REQUESTS PURSUANT TO 957 CMR 5.04 (Researchers, Payers, Providers, and Provider Organizations)
Please complete only if you are requesting Level 1 (de-identified) Case Mix.
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.

Not applicable, please see Section VI

VI. ALL OTHER REQUESTS - PURPOSE AND INTENDED USE

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

This research may identify policies that would result in lower health care costs for consumers by quantifying the surplus between insurance and hospital pairs and determining how it is divided. If the research suggests oligopolistic pricing practices which significantly disadvantage certain groups (i.e. MassHealth patients) or organizations (a hospital or insurance company), it could inform public policy and/or regulation to reduce costs for consumers.

2. **Attach** a brief (1-2 pages) description of your research methodology. (This description will not be posted on the internet.)
   Please find our research methodology attached to the end of this application

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).
   - ☒ 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.

VII. APPLICANT QUALIFICATIONS

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

Amanda Cook holds masters degrees in Economics from Vanderbilt and Purdue University. She is currently a PhD student in Economics at Purdue university. Her dissertation focuses on the effects of being uninsured on level of service and health outcomes using Maryland's inpatient data.
James Bland holds masters degrees in Economics from the University of Melbourne and Purdue University. He is currently a PhD student in Economics at Purdue university. His dissertation research involves experimental economics, in which human subjects make choices on the computer, and so he is familiar with confidentiality in using human subjects. Both researchers have extensive training in econometrics and data analysis and have completed HIPPA Certification at Purdue University.

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.)

VIII. DATA LINKAGE AND FURTHER DATA ABSTRACTION

Note: Data linkage involves combining CHIA data with other databases to create one extensive database for analysis. Data linkage is typically used to link multiple events or characteristics that refer to a single person in CHIA data within one database.

1. Do you intend to link or merge CHIA Data to other datasets?
   - ☒ Yes
   - ☐ No linkage or merger with any other database will occur

2. If yes, will the CHIA Data be linked or merged to other individual patient level data (e.g. disease registries, death data), individual provider level data (e.g., American Medical Association Physician Masterfile) , facility level (e.g., American Hospital Association data) or with aggregate data (e.g., Census data)? [check all that apply]

   ☒ Individual Patient Level Data

   What is the purpose of the linkage:

We would like to perform an aggregated merge (not a linkage) for a demographic comparison using service and case mix attributes between Case Mix and APCD data. Our goal is to compare charges by facility, aggregate demand attributes, diagnoses, and outcomes, taking advantage of DRGs in the case mix data.

   What databases are involved, who owns the data and which specific data elements will be used for linkage:

We would like to merge APCD and Case Mix data bases based on patient characteristics, hospital, and diagnosis.

   ☐ Individual Provider Level Data

   What is the purpose of the linkage:

   What databases are involved, who owns the data and which specific data elements will be used for linkage:

   ☒ Individual Facility Level Data

   What is the purpose of the linkage:

To create a complete picture of each hospital's characteristics in MA and to determine the extent to which two hospitals are substitutes.

What databases are involved, who owns the data and which specific data elements will be used for linkage:

Public use Chia databases. (Hospital cost reports, hospital charge use data, relative price reports, etc.)

☒ Aggregate Data

What is the purpose of the linkage:

The purpose of adding census data is to help control for factors unobserved in the hospital data alone, such as average income by zip code, total populations in geographic markets, etc.

What databases are involved, who owns the data and which specific data elements will be used for linkage:

American Community survey and other such data in the public domain. Data elements would include average income by zip code, total populations in geographic markets, etc.

3. If yes, for each proposed linkage above, please describe your method or selected algorithm (e.g., deterministic or probabilistic) for linking each dataset. If you intend to develop a unique algorithm, please describe how that algorithm will link each dataset.

For individuals, we will perform an aggregated merge by patient characteristics, services and hospital. For facility level data, we will match on facility characteristics, and for aggregate data linkages will match by city characteristics.

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

In merging the APCD and Case-Mix data sets we would be analyzing at a 'group' level: demographic comparisons, services, and case mix attributes. The other data we would add would be aggregate data, it would not provide more information about individuals.

5. If yes, and the data mentioned above is not in the public domain, please attach a letter of agreement or other appropriate documentation on restrictions of use from the data owner corroborating that they agree to have you initiate linkage of their data with CHIA data and include the data owner's website.

Not applicable.

IX. 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 would hope to publish our findings in an academic journal. Before publication, the paper may be available as a working paper. To protect patient confidentiality, we will only compute summary statistics, conditional means, and regression estimates over large groups (sample size larger than 11).

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.

Yes. They will read the working paper or the journal in which it is published. There typically is no fee to access a working paper, journals can be accessed through subscription or a fee per paper.

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. Will you be reselling the data?

☐ Yes
☒ No

If yes, in what format will you be reselling the data (e.g., as a standalone product, incorporated with a software product, with a subscription, etc.)?

none

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

X. 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:	None
Contact Person:	
Title:	
Address:	
Telephone Number:	
E-mail Address:	
Organization Website:	

8. Will the agent/contractor have access to the data at a location other than your location, your off-site server and/or your database?

   ☐ Yes
   ☐ No

   If yes, please provide information about the agent/contractor's data management practices, policies and procedures in your Data Management Plan.

9. Describe the tasks and products assigned to this agent or contractor for this project.

10. Describe the qualifications of this agent or contractor to perform such tasks or deliver such products.

11. Describe your oversight and monitoring of the activity and actions of this agent or subcontractor.

XIII. ASSURANCES

Applicants requesting and receiving data from CHIA pursuant to 957 CMR 5.00 ("Data Recipients") will be provided with data following the execution of a data use agreement that requires the Data Recipient to adhere to processes and procedures aimed at preventing unauthorized access, disclosure or use of data.

Data Recipients are further subject to the requirements and restrictions contained in applicable state and federal laws protecting privacy and data security, and will be required to adopt and implement policies and practices to protect CHIA data in a manner consistent with the requirements of the federal Health Insurance Portability and Accountability Act of 1996 (HIPAA).

Data Recipients must promptly notify CHIA of any unauthorized use or disclosure of CHIA data.

By my signature below, I attest to: (1) the accuracy of the information provided herein; (2) my organization's ability to meet CHIA's minimum data security requirements; and (3) my authority to bind the organization seeking CHIA data for the purposes described herein.

Signature:	
Printed Name:	Amanda Cook
Original Application Submission Date:	2/23/2015
Dates Application Revised:	

Research methodology:

We aim to quantify the bargaining power between hospitals and insurers over the price paid for services. All else held equal, insurers would rather pay a lower price for their customers' services at a hospital, and hospitals would rather receive a higher price for the services they provide to patients. As hospitals' and insurers' incentives to raise or lower prices are not aligned, we will model the determination of these prices as an unstructured bargaining problem, taking into account the surplus generated by an agreement on price, the consequences of not reaching an agreement, and the options available to the hospital and the insurer. In particular, if a hospital and insurer fail to reach an agreement, the hospital could charge higher prices for the out-of-network patients that would have been in-network had the pair reached an agreement. This would harm the insurer. However if the insurer is able to direct patients away from this hospital to avoid high prices, then the hospital will suffer a decline in admissions. Prices that result from such negotiations could therefore be function of, among other things:

- The services offered by the hospital (i.e. are there some services offered by this hospital that are not offered nearby?)
- The number of people covered by the insurer (i.e. if the insurer has more customers, the hospital will suffer a larger decline in admissions if they fail to agree on price)
- Whether there are nearby hospitals which the insurer could re-direct patients if they failed to agree on a price.

Our research aims to identify which of these are important factors.
Such negotiated prices mean that having insurance not only reduces the financial risk of bad health outcomes (i.e. substituting wealth in good times for wealth in bad times), but also, assuming that insurers pass on these savings, that those with insurance will expect to have lower charges for the same services, compared to those without insurance.
We will use the CHIA Case mix and APCD dataset, as well as publically available information, to first estimate a price multiplier associated with pairs of hospitals and insurers. That is, how much does a patient insured by insurer X pay for services at hospital Y, relative to the counter-factual cases where the patient was either (1) uninsured, or (2) insured by Medicaid? Answering this question will provide estimates of negotiated prices, relative to the price paid by a patient insured by Medicaid.
We will then use these estimates in a model of unstructured bargaining, to identify the important factors in determining prices. This model will be used to evaluate the outcomes in counter-factual scenarios such as universal government-sponsored healthcare, a market without insurance, price regulation, and the addition of new hospitals or insurers into the market.
Finally, we will compare the estimated prices that insurance companies pay the hospitals to premiums to determine if lower reimbursement rates to hospitals translate into lower premiums for patients and/or higher profits for insurance companies.