CHIA.
center for health information and analysis

Non-Government Application for Massachusetts All-Payer Claims Data [Exhibit A]

I. INSTRUCTIONS

This form is required for all Applicants, except Government Agencies as defined in 957 CMR 5.02, requesting protected health information. *All Applicants must also complete the Data Management Plan, attached to this Application. The Application and the Data Management Plan must be signed by an authorized signatory of the Organization. This Application and the Data Management Plan will be used by CHIA to determine whether the request meets the criteria for data release, pursuant to* 957 CMR 5.00. *Please complete the Application documents fully and accurately. Prior to receiving CHIA Data, the Organization must execute CHIA's Data Use Agreement. Applicants may wish to review that document prior to submitting this Application.*

Before completing this Application, please review the data request information on CHIA's website:

- Data Availability
- Fee Schedule
- Data Request Process

After reviewing the information on the website and this Application, please contact CHIA at apcd.data@state.ma.us if you have additional questions about how to complete this form.

All attachments must be uploaded to IRBNet with your Application. All Application documents can be found on the CHIA website in Word and in PDF format or on IRBNet in Word format. If you submit a PDF document, please also include a Word version in order to facilitate edits that may be needed.

Applications will not be reviewed until the Application and all supporting documents are complete and the required application fee is submitted. A Fee Remittance Form with instructions for submitting the application fee is available on the CHIA website and IRBNet. If you are requesting a fee waiver, a copy of the Fee Remittance Form and any supporting documentation must be uploaded to IRBNet.

II. FEE INFORMATION

1. Consult the most current Fee Schedule for All-Payer Claims Database data.

2. After reviewing the Fee Schedule, if you have any questions about the application or data fees, contact apcd.data@state.ma.us.

3. If you believe that you qualify for a fee waiver, complete and submit the Fee Remittance Form and attach it and all required supporting documentation with your application. Refer to the Fee Schedule (effective Feb 1, 2017) for fee waiver criteria.

4. Applications will not be reviewed until the application fee is received.

5. Data for approved Applications will not be released until the payment for the Data is received.

III. ORGANIZATION & INVESTIGATOR INFORMATION

Project Title:	Bundling in Competitive Markets with Adverse Selection
IRBNet Number:	1215636-1
Organization Requesting Data (Recipient):	Carnegie Mellon University
Organization Website:	www.cmu.edu
Authorized Signatory for Organization:	Linda Feuster Duffy
Title:	Associate VP of Sponsored Programs
E-Mail Address:	lfeuster@andrew.cmu.edu
Address, City/Town, State, Zip Code:	5000 Forbes Ave., Pittsburgh, PA 15213
Data Custodian: **(individual responsible for organizing, storing, and archiving Data)**	John Bigler
Title:	Director of Business Intelligence & Database Operations
E-Mail Address:	jsbigler@cmu.edu
Telephone Number:	412-268-5780
Address, City/Town, State, Zip Code:	5000 Forbes Ave., Pittsburgh, PA 15213
Primary Investigator (Applicant): **(individual responsible for the research team using the Data)**	Anh Nguyen
Title:	Assistant Professor of Economics
E-Mail Address:	Anhnguyen@cmu.edu
Telephone Number:	412-268-6191
Names of Co-Investigators:	Teck Yong Tan
E-Mail Addresses of Co-Investigators:	Teckyongtan@ntu.edu.sg

IV. PROJECT INFORMATION

1. What will be the use of the CHIA Data requested? [Check all that apply]

- ☐ Epidemiological
- ☐ Longitudinal Research
- ☐ Reference tool
- ☐ Surveillance
- ☐ Inclusion in a product
- ☐ Health planning/resource allocation
- ☐ Quality of care assessment
- ☑ Research studies
- ☐ Student research
- ☐ Other (describe in box below)
- ☐ Cost trends
- ☐ Rate setting
- ☐ Severity index tool
- ☐ Utilization review of resources

2. Provide an abstract or brief summary of the specific purpose and objectives of your Project. This description should include the research questions and/or hypotheses the project will attempt to address, or describe the intended product or report that will be derived from the requested data and how this product will be used. Include a brief summary of the pertinent literature with citations, if applicable.

This project seeks to understand how the government should regulate health insurance firms' bundling schemes. Although bundle discounts are ubiquitous in almost all product markets, when the market exhibits adverse selection, however, both firms' incentives to provide bundle discounts and their welfare implications have not been studied.

3. Has an Institutional Review Board (IRB) reviewed your Project?

☒ Yes [*If yes, a copy of the approval letter and protocol must be included with the Application package on IRBNet.*]
☐ No, this Project is not human subject research and does not require IRB review.

4. **Research Methodology**: Applicants must provide either the IRB protocol or a written description of the Project methodology (typically 1-2 pages), which should state the Project objectives and/or identify relevant research questions. This document must be included with the Application package on IRBNet and must provide sufficient detail to allow CHIA to understand how the Data will be used to meet objectives or address research questions.

V. PUBLIC INTEREST

1. Briefly explain why completing your Project is in the public interest. Use quantitative indicators of public health importance where possible, for example, numbers of deaths or incident cases; age-adjusted, age-specific, or crude rates; or years of potential life lost. *Uses that serve the public interest under CHIA regulations include, but are not limited to: health cost and utilization analysis to formulate public policy; studies that promote improvement in population health, health care quality or access; and health planning tied to evaluation or improvement of Massachusetts state government initiatives.*

Currently, ACA regulations forbid health insurance firms on the health insurance exchange to offer a bundle discount to households which have more members buying health insurance. Interestingly, this restriction was established not due to the potential competitive effects of bundling but as a simple way to ensure that age and tobacco use factors in premiums must be attributable to individuals.

In this project, we explore how health insurance premiums and consumer welfare will change when bundle discounts are allowed. We will also estimate the optimal level of bundle discount that the government can mandate through a tax rebate to household and its associated welfare improvement.

VI. DATA REQUESTED

The Massachusetts All-Payer Claims Database is comprised of medical, pharmacy, and dental claims and information from the member eligibility, provider, and product files that are collected from health insurance payers licensed to operate in the Commonwealth of Massachusetts. This information encompasses public and private payers as well as data from insured and self-insured plans. APCD data are refreshed and updated annually and made available to approved data users in Release Versions that contain five calendar years of data and three months of run-out. Data requests will be fulfilled using the most current Release Version. For more information about the most current APCD Release Version, including available years of data and a full list of elements in the release please refer to release layouts, data dictionaries and similar documentation included on CHIA's website.

Data requests are typically fulfilled on a one time basis, however; certain Projects may require future years of data that will become available in a subsequent release. Applicants who anticipate a need for future years of data may request to be considered for a subscription. Approved subscriptions will receive, upon request, the same data files and data elements included in the initial Release annually or as available. Please note that approved subscription request will be subject to the Data Use Agreement, will require payment of fees for additional Data, and subject to the limitation that the Data can be used only in support of the approved Project.

1. List years of data requested (only list years available in the current Release Version): 2012,2013,2014,2015,2016

2. Please indicate below whether this is a one-time request, or if the described Project will require a subscription.

☒ One-Time Request **OR** ☐ Subscription

3. Specify below the data files requested for this Project, and provide your justification for requesting *each* file.

☒ **Medical Claims**
Describe how your research objectives require Medical Claims data:
The Medical Claims data are used to construct annual medical expenditure for each individual, and the diagnosis code for each episode of care is used to construct individual's health status. These informations are then used to study the demand for health insurance.
☒ **Pharmacy Claims**
Describe how your research objectives require Pharmacy Claims data:
The Pharmacy Claims are used to construct annual medical expenditure for each individual, which is needed to estimate the demand for health insurance.
☐ **Dental Claims**
Describe how your research objectives require Dental Claims data:
☒ **Member Eligibility**
Describe how your research objectives require Member Eligibility data:
The Member Eligibility data include information related to the insurance product, in particular the start date and end date of enrollment and deductibles. These informations are needed to identify how consumers react to deductibles and coinsurance rates (i.e, moral hazard) , which explain their choice of health insurance plans.
☒ **Provider**
Describe how your research objectives require Provider data:
The Provider data allow us to disentangle consumers' preferences with respect to the financial characteristics of a health insurance plans and other non-financial characteristics such as provider network. If consumers have a preferred set of health care providers, the health insurance firm who has these providers in its network will have more market power.
☒ **Product**
Describe how your research objectives require Product data:
The Product data contains the start and end dates and the geographic market of each health insurance product that were available on the market. This is important to identify consumers' choice set, i.e, consumers who couldn't choose a particular plan because it was not available.

VII. DATA ENHANCEMENTS REQUESTED

State and federal privacy laws limit the release and use of Data to the minimum amount of data needed to accomplish a specific Project objective.

All-Payer Claims Database data is released in Limited Data Sets (LDS). All applicants receive the "Core" LDS, but may also request the data enhancements listed below for inclusion in their analyses. Requests for enhancements will be reviewed by CHIA to determine whether each represents the minimum data necessary to complete the specific Project objective.

For a full list of elements in the release (i.e., the core elements and additional elements), please refer to release layouts, data dictionaries and similar documentation included on CHIA's website.

1. Specify below which enhancements you are requesting in addition to the "Core" LDS, provide your justification for requesting each enhancement.

Geographic Subdivisions

The geographic subdivisions listed below are available for Massachusetts residents and providers only. Select *one* of the following options.

☑ 3-Digit Zip Code (standard)	☐ 5-Digit Zip Code***
***If requested, provide justification for requesting 5-Digit Zip Code. Refer to specifics in your methodology:	

Date Resolution

Select *one* option from the following options.

☒ Year (YYYY) (Standard)	☐ Month (YYYYMM) ***	☐ Day (YYYYMMDD) *** [for selected data elements only]
*** If requested, provide justification for requesting Month or Day. Refer to specifics in your methodology:		

National Provider Identifier (NPI)

Select *one* of the following options.

☒ Encrypted National Provider Identifier(s) (standard)	☐ Decrypted National Provider Identifier(s)***
*** If requested, provide justification for requesting decrypted National Provider Identifier(s). Refer to specifics in your methodology:	

VIII. MEDICAID (MASSHEALTH) DATA

1. Please indicate whether you are seeking Medicaid Data:

☐ Yes
☒ No

2. Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are ***directly connected to the administration of the Medicaid program***. If you are requesting MassHealth Data, please describe, in the space below, why your use of the Data meets this requirement. *Your description should focus on how the results of your project could be used by the Executive Office of Health and Human Services in connection with the administering the MassHealth program*. Requests for MassHealth 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 MassHealth program. CHIA cannot release MassHealth Data without approval from MassHealth. This may introduce significant delays in the receipt of MassHealth Data.

IX. DATA LINKAGE

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

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

2. If yes, please indicate below the types of data to which CHIA Data will be linked. [Check all that apply]
   ☐ Individual Patient Level Data (e.g. disease registries, death data)
   ☐ Individual Provider Level Data (e.g., American Medical Association Physician Masterfile)
   ☐ Individual Facility Level Data (e.g., American Hospital Association data)
   ☐ Aggregate Data (e.g., Census data)
   ☐ Other (please describe):

3. If yes, describe the dataset(s) to which the CHIA Data will be linked, indicate which CHIA Data elements will be linked and the purpose for each linkage.

4. 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 it will link each dataset.

5. If yes, attach or provide below a complete listing of the variables from all sources to be included in the final linked analytic file.

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

X. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Do you anticipate that the results of your analysis will be published or made publically available? If so, how do you intend to disseminate the results of the study (e.g.; publication in professional journal, poster presentation, newsletter, web page, seminar, conference, statistical tabulation)? Any and all publication of CHIA Data must comply with CHIA's cell size suppression policy, as set forth in the Data Use Agreement. Please explain how you will ensure that any publications ***will not disclose a cell less than 11***, and percentages or other mathematical formulas that result in the display of a cell less than 11.

We anticipate that the results of our analysis will be eventually published. The results will be presented in seminars and conferences and published in professional journal. We will only display summary statistics of samples that have at least 12 observations.

2. Describe your plans to use or otherwise disclose CHIA Data, or any Data derived or extracted from such Data, in any paper, report, website, statistical tabulation, seminar, or other setting that is not disseminated to the public.

We will not disclose CHIA Data or any Data derived or extracted from CHIA data in other settings that are not disseminated to the public

3. What will be the lowest geographical level of analysis of data you expect to present for publication or presentation (e.g., state level, city/town level, zip code level, etc.)? Will maps be presented? If so, what methods will be used to ensure that individuals cannot be identified?

Three digit zip code. Maps will be presented, but only at 3 digit zip code. If a zip code has less than 11 observation, it will be merged with neighboring zip codes in maps and tables.

4. Will you be using CHIA Data for consulting purposes?

☐ Yes
☒ No

5. Will you be selling standard report products using CHIA Data?

☐ Yes
☒ No

6. Will you be selling a software product using CHIA Data?

☐ Yes
☒ No

7. Will you be using CHIA Data as in input to develop a product (i.e., severity index took, risk adjustment tool, reference tool, etc.)

☐ Yes
☒ No

8. Will you be reselling CHIA Data in any format not noted above?

☐ Yes
☒ No

If yes, in what format will you be reselling CHIA Data?

9. If you have answered "yes" to questions 5, 6, 7 or 8, please describe the types of products, software, services, or tools.

10. If you have answered "yes" to questions 5, 6, 7 or 8, what is the fee you will charge for such products, software, services or tools?

XII. APPLICANT QUALIFICATIONS

1. Describe your previous experience using claims data. This question should be answered by the primary investigator and any co-investigators who will be using the Data.

We have not previously used claims data. However, we have previously worked on data that include sensitive information such birth date and addresses.

2. **Resumes/CVs**: When submitting your Application package on IRBNet, include résumés or curricula vitae of the principal investigator and co-investigators. (These attachments will not be posted on the internet.)

XIII. USE OF AGENTS AND/OR CONTRACTORS

By signing this Application, the Agency assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors. The Agency must have a written agreement with the agent of contractor limiting the use of CHIA Data to the use approved under this Application as well as the privacy and security standards set forth in the Data Use Agreement. CHIA Data may not be shared with any third party without prior written consent from CHIA, or an amendment to this Application. CHIA may audit any entity with access to CHIA Data.

Provide the following information for **all** agents and contractors who will have access to the CHIA Data. [*Add agents or contractors as needed.]*

AGENT/CONTRACTOR #1 INFORMATION	
Company Name:	Carnegie Mellon University will not be giving any agents and/or contractors access to CHIA data.
Company Website	
Contact Person:	
Title:	
E-mail Address:	
Address, City/Town, State, Zip Code:	
Telephone Number:	
Term of Contract:	

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

N/A

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

N/A

3. Will the agent or contractor have access to or store the CHIA Data at a location other than the Organization's location, off-site server and/or database?

☐ Yes
☐ No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

AGENT/CONTRACTOR #2 INFORMATION	
Company Name:	
Company Website:	Carnegie Mellon University will not be giving any agents and/or contractors access to CHIA data.
Contact Person:	
Title:	
E-mail Address:	
Address, City/Town, State, Zip Code:	
Telephone Number:	
Term of Contract:	

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

N/A

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

N/A

3. Will the agent or contractor have access to or store the CHIA Data at a location other than the Organization's location, off-site server and/or database?

☐ Yes
☐ No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

[INSERT A NEW SECTION FOR ADDITIONAL AGENTS/CONTRACTORS AS NEEDED]

IVX. ATTESTATION

By submitting this Application, the Organization attests that it is aware of its data use, privacy and security obligations imposed by state and federal law *and* confirms that it is compliant with such use, privacy and security standards. The Organization further agrees and understands that it is solely responsible for any breaches or unauthorized access, disclosure or use of CHIA Data, including, but not limited to, any breach or unauthorized access, disclosure or use by any third party to which it grants access.

Applicants approved to receive CHIA Data will be provided with Data following the payment of applicable fees and upon the execution of a Data Use Agreement requiring the Organization to adhere to processes and procedures designed to prevent unauthorized access, disclosure or use of data.

By my signature below, I attest: (1) to the accuracy of the information provided herein; (2) that the requested Data is the minimum necessary to accomplish the purposes described herein; (3) that the Organization will meet the data privacy and security requirements described in this Application and supporting documents, and will ensure that any third party with access to the Data meets the data use, privacy and security requirements; and (4) to my authority to bind the Organization.

Signature: (Authorized Signatory for Organization)	DocuSigned by:
Printed Name:	Linda Feuster Duffy
Title:	AVP for Sponsored Programs

Attachments

A completed Application must have the following documents attached to the Application or uploaded separately to IRBNet:

☐ 1. IRB approval letter and protocol (if applicable), or research methodology (if protocol is not attached)

☒ 2. Data Management Plan; including one for each agent or contractor that will have access to or store the CHIA Data at a location other than the Organization's location, off-site server and/or database

☒ 3. CVs of Investigators (upload to IRBnet)

APPLICATIONS WILL NOT BE REVIEWED UNTIL THEY ARE COMPLETE, INCLUDING ALL ATTACHMENTS.

[INSERT IRB approval letter and protocol, or research methodology]

Carnegie Mellon University will be submitting the IRB approval letter and protocol through IRB.net

Competitive Bundling in Markets with Adverse Selection

Anh Nguyen* and Teck Yong Tan†

January 2020

Research Question This project studies how family pricing should be regulated in a market with adverse selection problems such as the health insurance market. Part of the ACA regulations inadvertently forbid firms to use bundle discounts—i.e., the premium of a family plan is simply the sum of the individual premiums. We first show theoretically that this restriction softens competition between health insurance firms and the extent of this effect is positively dependent on consumers' preferences for different health insurance firms. In the empirical section, we would like to explore how alternative bundling regulations can reduce equilibrium premiums and improve welfare.

Research Methodology The purpose of the empirical section is two fold. First, we construct a framework to study and estimate the household's demand for health insurance that allows the household to choose different health insurance firms and health insurance plans for different household members. This is a novel contribution to the literature, which has often focused on the case in which all household members must choose the same health insurance plan. Second, we would like to use the empirical model to conduct counterfactual exercises that quantifies the magnitude of how alternative family bundle discount policies can improve welfare.

There are multiple reasons why the data from the Massachusetts All Payer Claim data are helpful for our analysis. First, as with any claim data, only insured individuals are observed. Since MA has the lowest uninsured rate in the country, we are less likely to suffer from sample selection issue of only observing insured individuals. Second, MA Health Connector has a stable number of participating health insurance firms, which alleviates concerns about the market not at equilibrium.

The medical claims and the pharmacy claims will be used to construct an estimate for the underlying health risks of each individual (similar to the approach used in [1]). Households'

*anhnguyen@cmu.edu. Tepper School of Business, Carnegie Mellon University
†teckyongtan@ntu.edu.sg. Department of Economics, Nanyang Technological University.

insurance choices (where a 'household' is identified as members who are in the same health insurance plan in at least one year) and the choices of health care providers help identify each household's risk preferences and preference for networks/health insurance firms. We are hoping to see switching into or out of the ACA pool by individuals who can purchase insurance from the employer sponsored insurance market as dependents and from the ACA over time, which will help disentangle individual household members' preferences from the household's aggregate preferences.

References

[1] Benjamin Handel."Adverse selection and inertia in health insurance markets: When nudging hurts." *American Economic Review*, 103.7 (2013): 2643-82. (document)

Addendum to APCD Application (IRBNet ID # 1215636-1)

Anh Nguyen*

February 12, 2020

Dear CHIA APCD Committee,

I would like to modify section IX (Data Linkage) of my data application (IRBNet ID # 1215636-1) submitted on 07/02/2019 as followed:

I intend to link CHIA Data to the zip code shape files downloaded from data.gov (https://catalog.data.gov/dataset/tiger-line-shapefile-2015-2010-nation-u-s-2010-census-5-digit-zip-code-tabulation-area-zcta5-na) to map information at the 3-digit zipcode level using the package 'tmap' in R. If a zip code has less than 11 observation, it will be merged with neighboring zip codes in maps. We would like to use these maps to illustrate the differences in demographics and medical spending between different geographical areas.

*Email: anhnguyen@columbia.edu

SPARCS IRB HUMAN SUBJECT RESEARCH

Printed on 18 Jan 2019 11:23:22 (irb version)

Print

Basic Information

1 * **Title of Study:**
Bundling in Competitive Markets with Adverse Selection

2 * **Brief description:**
Bundle discounts are ubiquitous in many product markets. However, preliminary analysis of the ACA health exchange shows that many insurers do not provide a bundle discount when a couple jointly purchase insurance. Such discounts are more often observed in states with lower average premium and fewer insurers. This does not resonate with the usual intuition of the bundling literature that firms almost always have an incentive to offer a bundle discount regardless of the degree of competition in the market. In this paper, we empirically show that adverse selection is the key reason why firms in some markets choose not to bundle. When a firm unilaterally offers a bundle discount, it risks attracting couples with worse health types from other firms. This adverse effect from bundle discounts is worsened when firms are more closely positioned in the product space. In the counterfactual exercise, we explore how a premium subsidy to reward couple purchase from the government could improve both consumer and producer surplus by mitigating adverse selection.

3 * **Principal investigator:**
Anh Nguyen

Title:
Assistant Professor of Economics

Department:
TEPPER FACULTY AFFAIRS

Study Team Members

For the purpose of this submission, study team members to be listed are CO-Is and Study Faculty Advisors.

We recognize that other team members may be involved with the work with the PI being responsible for appropriate training.

1 **Identify each Co-Investigator and Faculty Advisor. These should be individuals involved in the design, conduct or reporting of the research. Who should be included as a Co-I is at the discretion of the PI (see OHRP Guidance on "who are investigators?").**

Faculty Advisors must be listed for students serving as a PI.

All team members who interact with participants or who have access to identifiable research data, whether listed here or not, must complete CITI training on Human Subjects Research.

Name	Roles	Involved in Consent	E-mail	Phone	Title	Department
Teck Yong Tan	CO-I	no	teckyongtan@ntu.edu.sg			NANYANG TECHNOLOGICAL UNIVERSITY

2 **Identify any additional persons from external institutions who are also involved in the design, conduct, or reporting of the research, but who were not available in the preceding selector. NOTE: If this is a multi-site study requiring sIRB review at CMU, please include your CO-I's from all participating sites:**

First Name	Last Name	Email	Institution	SPARCS Account Requested

There are no items to display

3 * **Please briefly describe the qualifications and responsibilities of <u>each</u> study team member in regards to the research. Include the PI and all persons listed above. Please list a few sentences for each study team member, describing responsibilities and relevant expertise. (*This question must be answered to satisfy a regulatory requirement.*)**

> Examples:
>
> Andrew Carnegie will serve as PI on this project and will oversee all aspects of the research, from design through data analysis and publication of the results. Dr. Carnegie is a professor in the college of engineering who has 20 years of experience conducting research. His research interests include…
>
> Dr. Mellon will serve as a Co-I on this project. He is a physician from the Bayside Healthcare System who will oversee the research design, specific to the medical needs of the subject population. Dr. Mellon has been treating and researching this subject population for 30 years.
>
> Scotty Carnegie is a PhD candidate whose work will be overseen by the PI. He will assist with the data collection and analysis.

Anh Nguyen will serve as PI on this project and will oversea all aspects of the research, including data analysis and publication of the results. Dr. Nguyen is an assistant professor in Tepper School of Business, and her research Her research interest lies in the intersection of industrial organization and health economics, and development economics. Her current work focuses on the design of public health insurance programs in order to reduce the impact of adverse selection on social welfare.

Dr. Tan will serve as a Co-I on this project. He is an assistant professor at Nanyang Technological University in Singapore who will oversee the theoretical foundation for the empirical results.

Study Team Training

CITI Training for Study Team Members:

Name	Training	Uploaded Training Documentation
Teck Yong Tan	No training data to display	citiCompletionReport7216959.pdf(0.01)

CITI Training for Principal Investigator:

Course	Group	Stage	Completion Date	Expiration Date
Information Privacy Security (IPS) Course	Health Privacy	Basic Course	7/16/2018	
Social & Behavioral Research - Basic/Refresher	Social & Behavioral Research - Basic/Refresher	Basic Course	7/16/2018	7/15/2021

Training Documentation for Principal Investigator (only if PI is external to CMU):
CITI Completion for Teck Yong Tan(0.01)

Funding Sources

* **Is this funded research?**
(●) **Yes** ( ) No

Select SPARCS Funding Sources:

Project ID	Project Name	Project State	Project Type	PI	Sponsor Name
There are no items to display					

Select non-SPARCS Funding Sources:

	Funding Source (from list)	Funding Source (manual entry)	SPEX ID	Attachments	Is Internal / CMU Department
View	TEPPER SCHOOL OF BUSINESS AT CARNEGIE MELLON				yes

Review Type Requested

This choice will determine what type of review and approval your submission will receive as well as the type of questions which will follow this page. Requesting the correct type of review is important to avoid delays in the review. If you are unsure about which type of request is appropriate please contact the IRB Office at irb-review@andrew.cmu.edu.

1 * **What type of review are you requesting?**
Exempt – Cat 4

Protocol Description (Cat 4)

This category includes research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

1 **What will be collected or studied**? (*Please check all that apply)*

- [x] Existing Data
- [ ] Existing Documents
- [x] Existing Records
- [ ] Existing Pathological Specimens
- [ ] Existing Diagnostic Specimens

If the research involves other types of data, it is not eligible for this exemption. Do not proceed.

* **Is this information publically available?**
( ) Yes (●) **No**

* **Will information be recorded in such a manner that research subjects cannot be identified,**

directly or through identifiers linked to the subjects?
(●) **Yes** ( ) No

If the answers to the two preceding questions are both "no," the research is not eligible for this exemption. Do not proceed.

2 * **Provide, in lay terms, a summary of your proposed study:**
We would like to use claim data on medical spending and health insurance information from Massachusetts to study how the family structure affects health insurance selection in the household and how health insurers price premiums in response.

3 * **What is the purpose of the study (what is your research question) and how will the data collected be used?**
The purpose of the study is to see whether a government intervention that takes into account the family structure could improve the performance of the health insurance market.

The data from Massachusetts All Payer Claim Data will be used to identify members within the same household, their health insurance coverage, and their medical claim. We will then estimate the household's preference for health insurance as well as their expected medical costs.

Household identification is constructed using the link between the primary member of the health insurance plan (identified as the subscriber in the data), and his/her dependents in the same health insurance plan. No address information will be used.

Data/Documents/Records/Specimens (Cat 4)

1 * **From where will the data/documents/records/specimens be obtained?**
The data is provided by the Massachusetts Center for Health Information and Analysis.

2 * **Was the data collected with a consent form?**
( ) Yes (●) **No**

3 * **Was the data collected under a different IRB protocol?**
( ) Yes (●) **No**

4 * **Is de-identified information being provided to the investigator?**
( ) Yes (●) **No**

5 * **Will the information be recorded (documented on paper, electronically, digitally etc.) by the investigator?**
(●) **Yes** ( ) No

If yes, describe how the investigator will record data without identifiers:
We will assign a random numeric variable for each individual in lieu of the individual's ID assigned by the

original data from CHIA.

6 * **Is the information related to the provision of healthcare?**
(●) **Yes** ( ) No

If yes, does the data meet the HIPAA de-identification requirements?
( ) Yes (●) **No**

7 * **Describe the information being used in the research:**
Medical Claim Data:
1. Related to claim: year of submission, diagnosis codes, out-of-pocket payment (which includes co-payment, co-insurance, deductibles, etc), charge amount, referral provider ID (a numeric assigned by CHIA), type of visit (Primary care/referral).
2. Related to member: age of member at service, member relationship to primary health insurance subscriber, ID of health insurance product (to be matched with the product file, numeric variable assigned by CHIA), member's gender, member's zipcode (first 3 digits).
3. Related to health care provider: State and zip code.

Member Eligibility Data:
1. Information related to member (similar to what have been listed above).
2. Related to the insurance product: ID of health insurance product, whether the member is enrolled in dental/prescription drug/medical coverage, insurance type (HMO, PPO, etc), start date and end date of enrollment (in month/year), amount of deductibles for dental/prescription drug/medical spending, maximum coinsurance, monthly premium.

Pharmacy Claim Data:
1. Prescribing physician state and zip code, prescriber ID (to be matched with the medical claim data).
2. Out of pocket expense for each claim: co-payment amount, deductible amount, co-insurance.
3. Pharmacy zipcode and state
4. Diagnosis code for the claim, generic drug or branded.

Product Data: the start and end dates (in year) and the geographic market of each health insurance product being available on the market. Also includes individual and family deductible amounts.

Provider file: Provider ID to be linked with Medical claim data, payer-specific ID.

Data Security and Confidentiality (Cat 4)

1 * **Describe how you will protect anonymity and secure research records:**
We will not attempt to identify any individual based on the data. The data will be stored on a secured server.

Cooperating Institutions (Cat 4)

1. * **Is this research being done in cooperation with any institutions, individuals or organizations not affiliated with CMU?**
(●) **Yes** ( ) No

List Collaborating Institutions:

	Collaborators (from list)	Collaborators (Manual Entry)	Is CMU Overseeing Collaborator	Is Collaborator Overseeing CMU
View		Tan Teck Yong	no	no

2. * **Have you received IRB approval from another IRB for this study?**
No

Supporting Documents

1 **Attach supporting files:**

Document	Category	Date Modified	Document History

There are no items to display

Suggested attachments:

- Other study-related documents not attached on previous forms

Final Page

1 * **Does study have potential conflict of interest?**
○ Yes ◉ **No**

You have reached the end of the IRB submission form. Read the next steps carefully:

Click **Finish** to exit the form.

Important! To send the submission for review, the principal investigator must click **Submit** on the workspace.

Carnegie Mellon University

APPROVAL OF SUBMISSION

July 27, 2018

Type of Review:	Initial Study
Title of Study:	Bundling in Competitive Markets with Adverse Selection
Investigator: Study Team Members:	Anh Nguyen Teck Yong Tan
IRB ID:	STUDY2018_00000359: Bundling in Competitive Markets with Adverse Selection
Funding:	Name: TEPPER SCHOOL OF BUSINESS AT CARNEGIE MELLON; **********

The Carnegie Mellon University Institutional Review Board (IRB) has reviewed and granted **APPROVAL under as Exempt on 7/27/2018, in accordance with 45 CFR 46.101(b)(4)**.

This approval does not expire. However, if you wish to make modifications to this protocol, please contact the IRB regarding these changes prior to their implementation to ensure compliance with this designation.

The Investigator(s) listed above in conducting this protocol agree(s) to follow the recommendations of the IRB of any conditions to or changes in procedure subsequent to this review. In undertaking the execution of the protocol, the investigator(s) further agree(s) to abide by all CMU research policies including, but not limited to the policies on responsible conduct research and conflict of interest.

Sincerely,

John Zimmerman
IRB Chair

Carnegie Mellon University

APPROVAL OF SUBMISSION

January 15, 2019

Type of Review:	Modification
Title of Study:	Bundling in Competitive Markets with Adverse Selection
Investigator: Study Team Members:	Anh Nguyen Teck Yong Tan
IRB ID:	MOD00001985: Modification #1 for Study STUDY2018_00000359
Funding:	Name: TEPPER SCHOOL OF BUSINESS AT CARNEGIE MELLON; **********

This is to notify you that your **modification request** was approved on **1/15/2019**.

Approval for this study does not expire.

Please be reminded that if additional changes are to be made, those changes will need to be reviewed prior to implementation.

Sincerely,

John Zimmerman
IRB Chair