Wine, Alcohol, and Cardiovascular Health: Revisiting the Health Benefits of Wine in the Framingham Heart Study.

Michael Darden and Douglas Nelson

May 28th, 2015
Outline

Background

Our Approach

Framingham Heart Study
  Descriptive Analysis

Application - Smoking
Does Everything We Consume Cause Cancer?
Does Everything We Consume Cause Cancer?

Everything we eat both causes and prevents cancer

- Wine
- Tomatoes
- Tea
- Milk
- Eggs
- Corn
- Coffee
- Butter
- Beef

Relative risk of cancer

Protects against cancer Causes cancer

SOURCE: Schoenfeld and Ioannidis, American Journal of Clinical Nutrition
Assessing the long-term health effects of alcohol consumption is difficult because:

1. Many major health outcomes occur far in the future.
2. Ideally the difference in major health outcomes should be calculated from observed health differences following random assignment of lifetime drinking behavior.

▶ In reality, researchers must use observational data.
Assessing the long-term health effects of alcohol consumption is difficult because:

1. Many *major* health outcomes occur far in the future.
Assessing the long-term health effects of alcohol consumption is difficult because:

1. Many *major* health outcomes occur far in the future.
2. Ideally the difference in major health outcomes should be calculated from observed health differences following random assignment of lifetime drinking behavior.
Why is it difficult?

Assessing the long-term health effects of alcohol consumption is difficult because:

1. Many *major* health outcomes occur far in the future.
2. Ideally the difference in major health outcomes should be calculated from observed health differences following random assignment of lifetime drinking behavior.
   - In reality, researchers must use observational data.
What is typically done?

1. Difference in Means
   - Simple and transparent
   - Ignores potential heterogeneity

For example:
- Doll et al. 2004, *British Medical Journal* study the effect of smoking on expected longevity
- Confounding factors “are unlikely to have influenced greatly the absolute difference between the overall mortality rates of cigarette smokers and lifelong non-smokers” - Doll et al. 2004, *British Medical Journal*
What is typically done?

1. Difference in Means
   - Simple and transparent
   - Ignores potential heterogeneity
For example:
   - Doll et al. 2004, *British Medical Journal* study the effect of smoking on expected longevity
1. Difference in Means
   - Simple and transparent
   - Ignores potential heterogeneity

For example:
   - Doll et al. 2004, *British Medical Journal* study the effect of smoking on expected longevity
   - Confounding factors “are unlikely to have influenced greatly the absolute difference between the overall mortality rates of cigarette smokers and lifelong non-smokers” - Doll et al. 2004, *British Medical Journal*
2. “Controlling” for observed heterogeneity.
   ▶ Conditional estimates of health impact.
   ▶ Biased if unobserved factors influence both lifestyle behavior and health outcomes
What is typically done?

2. “Controlling” for observed heterogeneity.
   - Conditional estimates of health impact.
   - Biased if unobserved factors influence both lifestyle behavior and health outcomes

For example:
   - Djoussee et al. 2004, *Stroke* study the effect of alcohol consumption on the risk of Ischemic Stroke
2. “Controlling” for observed heterogeneity.
   ▶ Conditional estimates of health impact.
   ▶ Biased if unobserved factors influence both lifestyle behavior and health outcomes

For example:
   ▶ Djoussee et al. 2004, Stroke study the effect of alcohol consumption on the risk of Ischemic Stroke
   ▶ “The full model controlled for age, diabetes mellitus(yes/no), smoking categories, and body mass index.”
Our Approach

Consider the following question:

▶ “How many years of expected longevity are lost from heavy alcohol consumption?”

We argue this question is difficult to answer because:

1. Lifetime drinking patterns reflect choices of individuals who differ in observed and unobserved ways. That is, alcohol consumption is not random.

2. Drinkers and nondrinkers may differ in observed and unobserved ways that contribute to mortality. That is, confounding factors influence health outcomes.

Because random assignment is not possible, we argue that important qualities of observational data should include:

▶ A long, longitudinal study design.

▶ Frequent interviews/exams.

▶ Rich observable heterogeneity.
Our Approach

Consider the following question:

▶ “How many years of expected longevity are lost from heavy alcohol consumption?”

We argue this question is difficult to answer because:

1. Lifetime drinking patterns reflect choices of individuals who differ in observed and unobserved ways. That is, alcohol consumption is not random.
2. Drinkers and nondrinkers may differ in observed and unobserved ways that contribute to mortality. That is, confounding factors influence health outcomes.

Because random assignment is not possible, we argue that important qualities of observational data should include:

▶ A long, longitudinal study design.
▶ Frequent interviews/exams.
▶ Rich observable heterogeneity.
Our Approach

Consider the following question:

▶ “How many years of expected longevity are lost from heavy alcohol consumption?”

We argue this question is difficult to answer because:

1. Lifetime drinking patterns reflect choices of individuals who differ in observed and unobserved ways. That is, alcohol consumption is not random.

2. Drinkers and nondrinkers may differ in observed and unobserved ways that contribute to mortality. That is, confounding factors influence health outcomes.

Because random assignment is not possible, we argue that important qualities of observational data should include:

▶ A long, longitudinal study design.
▶ Frequent interviews/exams.
▶ Rich observable heterogeneity.
Our Approach

Consider the following question:

▶ “How many years of expected longevity are lost from heavy alcohol consumption?”

We argue this question is difficult to answer because:

1. Lifetime drinking patterns reflect choices of individuals who differ in observed and unobserved ways. That is, alcohol consumption is not random.

2. Drinkers and nondrinkers may differ in observed and unobserved ways that contribute to mortality. That is, confounding factors influence health outcomes.
Our Approach

Consider the following question:

▶ “How many years of expected longevity are lost from heavy alcohol consumption?”

We argue this question is difficult to answer because:

1. Lifetime drinking patterns reflect choices of individuals who differ in observed and unobserved ways. That is, alcohol consumption is not random.

2. Drinkers and nondrinkers may differ in observed and unobserved ways that contribute to mortality. That is, confounding factors influence health outcomes.

Because random assignment is not possible, we argue that important qualities of observational data should include:

▶ A long, longitudinal study design.
▶ Frequent interviews/exams.
▶ Rich observable heterogeneity
Outline

Background

Our Approach

Framingham Heart Study
  Descriptive Analysis

Application - Smoking
The Framingham Heart Study

- Begun in 1948 to "identify the common factors that contribute to cardiovascular disease".
- Sampled 5209 individuals from Framingham, MA and followed them for over 50 years.
- Detailed health information. Fairly consistent alcohol information.
- Ages range from 30 to 63 at the initial exam (between 1948 and 1953).

The Offspring Cohort:
- Offspring of the Original Cohort began followed in 1971.
- Sample of 5124 people between ages of 13 and 62.
- Health examinations/interviews every 2-8 years.
The Framingham Heart Study

The Original Cohort

- Begun in 1948 to “identify the common factors that contribute to cardiovascular disease”.
- Sampled 5209 individuals from Framingham, MA and followed them for over 50 years.
- Detailed health information. Fairly consistent alcohol information.
- Ages range from 30 to 63 at the initial exam (between 1948 and 1953).

The Offspring Cohort:

- Offspring of the Original Cohort began followed in 1971
- Sample of 5124 people between ages of 13 and 62
- Health examinations/interviews every 2-8 years.
Beer, wine, and spirits (drinks per week) was recorded at exams:

- Original Cohort: 12-15, 17-23
- Offspring Cohort: 1-7
Beer, wine, and spirits (drinks per week) was recorded at exams:

- Original Cohort: 12-15, 17-23
- Offspring Cohort: 1-7

In addition to a variety of biomarkers measured in each exam, we have dates for the following health events:

- Coronary Heart Disease, Myocardial Infarction, Angina Pectoris, Coronary Insufficiency, Cerebrovascular Accident, Atherothrombotic Infarction, Transient Ischemic Attack, Cardiovascular Disease, Intermittant Claudication, Congestive Heart Failure, etc.
- A wide variety of cancers.
- Death
Sample Construction

Beer, wine, and spirits (drinks per week) was recorded at exams:

- Original Cohort: 12-15, 17-23
- Offspring Cohort: 1-7

In addition to a variety of biomarkers measured in each exam, we have dates for the following health events:

- Coronary Heart Disease, Myocardial Infarction, Angina Pectoris, Coronary Insufficiency, Cerebrovascular Accident, Atherothrombotic Infarction, Transient Ischemic Attack, Cardiovascular Disease, Intermittent Claudication, Congestive Heart Failure, etc.
- A wide variety of cancers.
- Death

We pool data from both cohorts and, after restricting our sample such that a person enters at least twice, we have 7,454 unique individuals and 46,696 person/exam observations.
### Table: Summary Statistics - Exogenous Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Original Cohort</th>
<th></th>
<th>Offspring Cohort</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. Dev</td>
<td>Mean</td>
<td>St. Dev</td>
</tr>
<tr>
<td>Female</td>
<td>0.579</td>
<td>0.494</td>
<td>0.518</td>
<td>0.500</td>
</tr>
<tr>
<td>College</td>
<td>0.302</td>
<td>0.459</td>
<td>0.480</td>
<td>0.500</td>
</tr>
<tr>
<td>Age</td>
<td>71.827</td>
<td>8.432</td>
<td>52.974</td>
<td>11.498</td>
</tr>
<tr>
<td>Body Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI &lt; 20</td>
<td>0.040</td>
<td>0.195</td>
<td>0.042</td>
<td>0.201</td>
</tr>
<tr>
<td>BMI [20, 25)</td>
<td>0.332</td>
<td>0.471</td>
<td>0.333</td>
<td>0.471</td>
</tr>
<tr>
<td>BMI [25, 29)</td>
<td>0.399</td>
<td>0.490</td>
<td>0.398</td>
<td>0.489</td>
</tr>
<tr>
<td>BMI [30, 35)</td>
<td>0.133</td>
<td>0.340</td>
<td>0.161</td>
<td>0.368</td>
</tr>
<tr>
<td>BMI &gt; 35</td>
<td>0.036</td>
<td>0.187</td>
<td>0.066</td>
<td>0.248</td>
</tr>
<tr>
<td>BMI Missing</td>
<td>0.060</td>
<td>0.237</td>
<td>0.005</td>
<td>0.071</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>140.460</td>
<td>20.779</td>
<td>125.711</td>
<td>18.266</td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>224.484</td>
<td>42.555</td>
<td>205.559</td>
<td>39.146</td>
</tr>
<tr>
<td>Missing</td>
<td>0.399</td>
<td>0.490</td>
<td>0.026</td>
<td>0.159</td>
</tr>
<tr>
<td>Currently Smoke</td>
<td>0.187</td>
<td>0.390</td>
<td>0.240</td>
<td>0.427</td>
</tr>
<tr>
<td>Years Smoking</td>
<td>17.084</td>
<td>18.505</td>
<td>14.868</td>
<td>14.602</td>
</tr>
</tbody>
</table>
**Table: Summary Statistics - Alcohol**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Original Cohort</th>
<th>Offspring Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean St. Dev</td>
<td>Mean St. Dev</td>
</tr>
<tr>
<td>Drinks Wine</td>
<td>0.324 0.468</td>
<td>0.447 0.497</td>
</tr>
<tr>
<td>Glasses/week</td>
<td>3.733 4.954</td>
<td>4.032 4.584</td>
</tr>
<tr>
<td>Drinks Beer</td>
<td>0.202 0.401</td>
<td>0.341 0.474</td>
</tr>
<tr>
<td>Drinks Spirits</td>
<td>0.432 0.495</td>
<td>0.357 0.479</td>
</tr>
<tr>
<td>Beers/week</td>
<td>6.512 7.691</td>
<td>5.250 7.207</td>
</tr>
</tbody>
</table>
Drinks by Age

Drinks Alcohol by Type and Age

- Drinks Wine
- Drinks Beer
- Drinks Spirits
Drinks/Week by Age

Drinks per week by Type and Age
Original Cohort Health Events

Health Events by Observation - Original Cohort

- Cardiovascular Disease
- Cancer
Health Events by Observation - Offspring Cohort

Observation

Cardiovascular Disease Cancer

Percent

Observation

Cardiovascular Disease
Cancer

0 1 2 3 4 5 6 7

0 5 10 15

AWEA, Wine and Health, 05/28/2015
Original Cohort Survival

Kaplan-Meier Survival Curves - Original Cohort

Drinker Never Drinker

AWEA, Wine and Health, 05/28/2015
Kaplan-Meier Survival Curves - Offspring Cohort

- Drinker
- Never Drinker

Age
Dynamic Panel Data Model. $t$ subscript represents FHS exam time:

- **Dependent Variables:**
  1. Drinks wine at $t$
  2. Glasses per week | drinking wine at $t$
  3. Drinks beer at $t$
  4. Beers per week | drinking beer at $t$
  5. Drinks Spirits at $t$
  6. Cocktails per week | drinking spirits at $t$
  7. Cardiovascular Disease at $t$
  8. Cancer at $t$
  9. Death at $t$
Econometric Framework

\[\text{Wit} = w(A_{it} - 1, H_{it} - 1, X_{it}, P_{it}, u_{w_{it}}, k_{w_{it}})\]

\[H_{it} = h(A_{it}, H_{it} - 1, X_{it}, h_{it}, k_{h_{it}})\]

\[M_{it} = h(A_{it}, H_{it}, X_{it}, m_{it}, k_{m_{it}})\]
Econometric Framework

\[ W_{it} = w(A_{it-1}, H_{it-1}, X_{it}, P_{it}, u^w_{it}, \epsilon^w_{it}) \]
Econometric Framework

\[ W_{it} = w(A_{it-1}, H_{it-1}, X_{it}, P_{it}, u_{it}^w, \epsilon_{it}^w) \]

\[ H_{it} = h(A_{it}, H_{it-1}, X_{it}, \mu_{it}^h, \epsilon_{it}^h) \]
Econometric Framework

\[ W_{it} = w(A_{it-1}, H_{it-1}, X_{it}, P_{it}, u_{it}^w, \epsilon_{it}^w) \]

\[ H_{it} = h(A_{it}, H_{it-1}, X_{it}, \mu_{it}^h, \epsilon_{it}^h) \]

\[ M_{it} = h(A_{it}, H_{it}, X_{it}, \mu_{it}^m, \epsilon_{it}^m) \]
Error Structure

The error structure for equation $j$:

$$\mu_i + \nu_t + \epsilon_{it}$$
Error Structure

The error structure for equation $j$:

$$\mu_i + \nu_t + \epsilon_{it}$$

The distribution of both time-invariant and time-varying unobserved heterogeneity terms, $\mu_i$ and $\nu_t$, are approximated discretely. We estimate a step function for both:

- $\mu_i$ takes $A$ points: $(\mu_1, \ldots, \mu_A)$.
- $\nu_t$ takes $B$ points: $(\nu_1, \ldots, \nu_B)$.

Subject to normalizations, we estimate these points for each equation as well as the associated probabilities.
Model Setup

- The model captures alcohol and health outcomes for up to 11 periods.
- Each period of the model is two years in length.

\[
W_t \quad H_t \quad M_t
\]

\[
W_{t-1} \quad H_{t-1} \quad M_{t-1}
\]
**Table: Selected Parameter Estimates**

<table>
<thead>
<tr>
<th>Wine Quartile Relative to No Wine Consumption</th>
<th>CVD</th>
<th>Cancer</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>-0.215*</td>
<td>-0.067</td>
<td>-0.228</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.149)</td>
<td>(0.175)</td>
</tr>
<tr>
<td>2nd</td>
<td>-0.213*</td>
<td>0.124</td>
<td>-0.235</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.146)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>3rd</td>
<td>-0.108</td>
<td>0.009</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.099)</td>
<td>(0.129)</td>
<td>(0.133)</td>
</tr>
<tr>
<td>4th</td>
<td>-0.045</td>
<td>-0.063</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.160)</td>
<td>(0.158)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alcohol Quartile Relative to No Alcohol Consumption</th>
<th>CVD</th>
<th>Cancer</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>-0.248**</td>
<td>-0.049</td>
<td>-0.295**</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.094)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>2nd</td>
<td>-0.212**</td>
<td>-0.114</td>
<td>-0.278**</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.104)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>3rd</td>
<td>-0.238**</td>
<td>-0.143</td>
<td>-0.339**</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.105)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>4th</td>
<td>-0.193**</td>
<td>0.147</td>
<td>-0.214**</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.090)</td>
<td>(0.098)</td>
</tr>
</tbody>
</table>
Simulation Exercise

To simulate our model we:

1. Expand each person’s observation by 50
Simulation Exercise

To simulate our model we:

1. Expand each person’s observation by 50
2. Endow each with random draws from all error terms
Simulation Exercise

To simulate our model we:

1. Expand each person’s observation by 50
2. Endow each with random draws from all error terms
3. Simulate Behavior *while updating state vector.*
   - In some cases we impose alcohol consumption and simulate health outcomes.
4. Collapse results by person, draw, and time.
Kaplan-Meier Survival Curves: Data

Source: Authors’ Calculations from the Framingham Heart Study
<table>
<thead>
<tr>
<th>Smoking Behavior</th>
<th>Mean Age</th>
<th>Mean Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never Smoker</td>
<td>75.68 (8.81)</td>
<td>-</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>66.75 (10.10)</td>
<td>-8.92</td>
</tr>
</tbody>
</table>
Survival Curves from the Framingham Heart Study

Kaplan-Meier Survival Curves: Data

Source: Authors’ Calculations from the Framingham Heart Study
Simulated Survival Curves by Lifetime Smoking Behavior

Kaplan-Meier Survival Curves: Simulation

Source: Authors’ Calculations
Simulated Mean Age of Death by Smoking Behavior

<table>
<thead>
<tr>
<th></th>
<th>Mean Age</th>
<th>Mean Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never Smoker</td>
<td>75.60</td>
<td>-</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>71.33</td>
<td>-4.27</td>
</tr>
</tbody>
</table>
Thank you!