Wearable Technologies and Health Behaviors: New Data and New Methods to Understand Population Health

Benjamin Handel and Johnston Kolstad

Presented by: Dongning Shao

September 2019

Outline

Introduction: Background

- Two contradictory facts:
  - Huge Health care expenditure: in 2014, US spent 3 trillion dollars ($9523 per capita) or 17.5% GDP in Health care, more than any other countries
  - Disproportionate life expectancy: life expectancy (and other measures of health condition) is worse than other OECD counterparts
- Poor health related behaviours are crucial causes of this gap
- Evidences shows up to **one-third** of the death in US resulted from poor health related behaviours(Loewenstein et al., 2007)
- It is crucial for economists and policy makers to find out the cause of the gap as well as possible solution

Introduction: Motivation

- One possible explanation: private return of better behaviour is too small relative to the social cost (externality!), but evidence (Jackevicius et al., 2002) says this is insufficient
- Need to address "**internality**" in addition to externality
- One way to deal with this internality is to help individual
  1. identify the health impact of various of behaviours
  2. set up a plan to carry out good behaviours
  3. monitor and support to stick to the plan

Research Question

- Milkman et al. (2011) shows individuals who plan to improve behaviour and stick to the plan see better outcome in a small-scale experiment setting
- **Research Question**: Can we externalize the result in a large-scale experiment?

Contribution

- This paper uses the data collected by wearable technology and other IT tools in a context of large scale RCT in a large employer
- The result shows statistically significant yet economically small effect of using wearable technology to plan and improve health related behaviours
- The treatment heterogeneity is unlikely to be substantial. Therefore, further personalization, in this context, is futile

Empirical Strategy

- A large scale RCT between Oct 2015 and May 2016, subsidized a large portion (75%) of the retail price to encourage adoption of a wearable device
- The treatment gives the user access to a web-based tool which allows users to **upload data, monitor performance and make individual plans to improve behaviours**
- 20211 individuals received subsidy, 17276 purchased the device and of this group 14911 connected to the web-based tool to be included in the study
- 75% of the final study group were invited to enroll a health plan (the treatment group), among the treatment group, 27% joined sleep improve plan and 20% joined exercise improve plan; remaining 3600 individuals are the control group

Empirical Strategy

- **Treatment group**: access to upload data, monitor the performance and use planning tools
  Control group: only have access to upload data
- The simple difference between two groups is Intent To Treat (ITT) impact
- ITT impact causes some trouble because not all control group users enrolled in the plan while some others were actually able to access a plan (to use planning tools).
- To properly identify the effect of plans, authors used treatment assignment as an IV

Result: Population sleep statistics

A simple model with individual fixed effect and daily fixed effect is presented as a baseline,

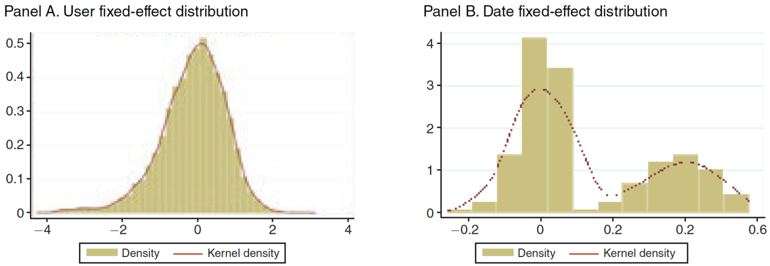

FIGURE 1. DISTRIBUTION OF INDIVIDUAL AND DAY FIXED EFFECTS FOR SLEEP HOURS

Result: Population sleep statistics

- Population average sleep hour is 6.3 hr/night.
- Panel A: a small share is sleeping consistently at recommended level (7-9 hrs) and a small number of users consistently sleep very little.
- Panel B: for individual-wise variation, the daily effect shows a large share of days with no significant difference, but there is also a mass above the average level. This captures the pattern that most people sleep roughly the same hours on weekdays with some time increase on weekends.

Result: Average Impact

Table 1—Mean Outcome for Treatment and Control with Different Sample Restrictions

	Treatment mean	Control mean	Diff	IV
Panel A. Base results				
Steps per day	6,040.32	5,988.47	51.85	154.8
Cardio score	1,069.00	1,056.38	12.62	37.55
Sleep time (hours)	6.30	6.28	0.02	0.0627
Sleep recovery quality	43.02	42.92	0.10	0.300
Hours worn per day	16.15	16.02	0.13	0.349
Panel B. First month of treatment				
Steps per day	5,390.95	5,322.77	68.18	230.3
Cardio score	953.44	928.34	25.10	84.19
Sleep time (hours)	6.35	6.32	0.03	0.114
Sleep recovery quality	43.92	43.83	0.08	0.301
Hours worn per day	16.42	16.28	0.13	0.386
Panel C. Low groups				
Steps per day	3,885.81	3,852.78	33.03	103.0
Cardio score	588.84	614.39	−25.55	−80.00
Sleep time (hours)	4.92	4.99	−0.07	−0.252
Sleep recovery quality	39.25	37.87	1.37	4.415
Hours worn per day	15.21	14.67	0.54	1.573

Result: Average Impact

- Panel A: the IIT estimate is statistically significant but economically small, IV estimate is larger but remains small in absolute magnitude; hours worn are similar
- Panel B: same approach, sample limited to the period from Jan 15 to Feb 15 (an early stage) considering the higher engagement and lower differential attrition influence, similar result
- Panel C: same approach, sample limited to those with a low baseline level. Cardio score and sleep hours actually decline with treatment, sleep quality seems to improve and hours of worn seems to be higher

Result: Heterogeneous Treatment Effect

- Authors also looked at the heterogeneous treatment effect in fear that the low ATE may mask the the sub group effect that may be substantial
- Implemented a machine learning approach developed by Athey and Imbens (2015)
- The result suggests little heterogeneous treatment effect, as a function of observables, presented in the impact of accessing a plan
- Based on the result, author believes further personalize planning tools are unlikely to make much improvement

Conclusion

There are some limitation in this research:

- This research is unable to identify the impact of simply being able to see personal data on behaviour
- The tool involved and setting applied is quite specific, hence one can not draw similar conclusion on other approaches
- Authors cannot rule out other unmeasured factors may be important on the heterogeneous treatment effect due to the features of the machine learning technique

Despite these limitations, wearable data has important potential for measurement of economic data

Reference

Athey, S. and Imbens, G. W. (2015). Machine learning methods for estimating heterogeneous causal effects. *stat*, 1050(5):1–26.

Jackevicius, C. A., Mamdani, M., and Tu, J. V. (2002). Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *Jama*, 288(4):462–467.

Loewenstein, G., Brennan, T., and Volpp, K. G. (2007). Asymmetric paternalism to improve health behaviors. *Jama*, 298(20):2415–2417.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. *Proceedings of the National Academy of Sciences*, 108(26):10415–10420.