

CAREER-DML: A Framework for Causal Inference on Career Trajectories

Rodolf Mikel Ghannam Neto
Independent Researcher

February 20, 2026

Abstract

This paper introduces CAREER-DML, a framework for estimating heterogeneous causal effects of career transitions using sequential data. We leverage Double/Debiased Machine Learning (DML) combined with recurrent neural network (RNN) embeddings to model high-dimensional career histories and estimate the causal impact of transitioning to an AI-exposed occupation. Our primary contribution is twofold. First, we identify and document a **Sequential Embedding Ordering Phenomenon** in this domain: causally-regularized embeddings (Variational Information Bottleneck), which are theoretically appealing, consistently exhibit higher estimation bias than simpler predictive embeddings. This empirical finding challenges the direct application of certain causal representation learning methods to socio-economic trajectories and calls for further investigation. Second, we characterize a **Signal-to-Noise Frontier**, demonstrating that while our framework substantially outperforms classical methods like Heckman selection models (reducing estimation bias by an order of magnitude, from 945% to under 60%), detecting realistic, small-magnitude treatment effects (e.g., an 8% wage premium) requires sample sizes ($N \geq 1,000$) beyond those typical in survey data, motivating the use of large-scale administrative registers. We validate our framework on a **semi-synthetic data generating process (DGP)**, which is carefully calibrated with real-world US labor market data, and provide a comprehensive open-source implementation.

Keywords: Causal Inference, Double Machine Learning, Representation Learning, Labor Economics, AI Adoption.

JEL Codes: C21, C55, J24, J31.

* Author's contact: rodolf@cical.com.br. This paper is submitted as supplementary material for the author's PhD application to the Copenhagen Business School. The codebase is available at <https://github.com/RodolfGhannam/CAREER-DML>.

1 Introduction

The proliferation of Artificial Intelligence (AI) is projected to fundamentally reshape labor markets, creating both opportunities and disruptions (Acemoglu and Restrepo, 2019). This transformation is no longer merely projected; it is underway. According to McKinsey, 78% of organizations now use AI in at least one business function, up from 55% in 2023 (McKinsey & Company, 2025). In the Nordic countries, individual AI uptake has surpassed 35%, with 34.7% of the Danish workforce exposed to generative AI in tasks representing 20% or more of their work hours (OECD, 2026, 2024). The pace and scale of this adoption make the causal question—what happens to workers who transition into AI-exposed occupations—both urgent and empirically tractable. A central question for policymakers and individuals is to understand the causal effect of adapting to this new reality, for instance, by transitioning into an occupation with high AI exposure. Estimating this effect is challenging due to high-dimensional confounding; an individual’s career history—a sequence of occupations, wages, and educational choices—is a powerful determinant of both their future career choices and their earnings potential (Neal, 1999).

Traditional econometric methods struggle with such high-dimensional, sequential data. This paper proposes CAREER-DML, a framework that combines the flexibility of deep learning for representation learning with the robustness of Double/Debiased Machine Learning (DML) for causal inference (Chernozhukov et al., 2018). The core idea is to use a Recurrent Neural Network (RNN) to embed an individual’s entire career history into a fixed-length vector, z , which then serves as a high-dimensional control variable in a DML estimation.

1.1 Dialogue with Structural Labor Economics

While our approach is rooted in the reduced-form tradition of causal ML, it builds an explicit bridge to the **structural econometrics tradition** of the labor market. The learned embedding z can be interpreted as a rich, non-parametric approximation of key latent variables in structural models:

- **Human Capital Stock:** In the tradition of Ben-Porath (1967), z can be seen as an empirical measure of an individual’s human capital stock, capturing not just years of experience but its quality, relevance, and depreciation as learned from the career sequence.
- **Generalized Mincer Experience:** Our model, $Y = \theta T + g(z) + \epsilon$, acts as a non-parametric extension of the Mincer (1974) earnings equation. The function $g(z)$ replaces the simple quadratic in experience with a flexible function of a much richer representation of a worker’s entire career path.
- **Latent Task Space:** Following the task-based framework of Autor et al. (2003), the embedding z implicitly learns a latent “task space” by observing sequences of occupational

transitions, allowing us to analyze how technology shocks affect workers positioned differently within this space.

- **Empirical Latent Type:** In discrete choice dynamic programming models (e.g., [Keane and Wolpin \(1997\)](#)), unobserved heterogeneity is often modeled via a fixed number of latent "types." The embedding z can be viewed as a continuous, high-dimensional, and empirically-estimated measure of this latent type.

By estimating these objects flexibly, CAREER-DML provides inputs that can be used to calibrate, test, or enrich structural models.

1.2 Core Empirical Findings

We explore three classes of RNN embeddings: **Predictive**, **Causal (VIB)** ([Veitch et al., 2020](#)), and **Adversarial**. Our analysis, conducted in a carefully calibrated semi-synthetic environment, yields two primary findings.

First, we document a **Sequential Embedding Ordering Phenomenon**: the causally-motivated VIB embeddings consistently produce *more* biased estimates of the treatment effect than simpler predictive embeddings. This empirical, and perhaps counterintuitive, ordering suggests that the information bottleneck, while effective in other domains, may discard causally crucial information present in the temporal ordering of career data. This highlights the need for domain-specific adaptation of causal representation learning methods.

Second, we characterize a **Signal-to-Noise Frontier**. While all ML-based methods dramatically outperform the classical Heckman model ([Heckman, 1979](#)), we find that at realistic effect sizes (an 8% wage premium), the incremental gain from sequential embeddings over simpler static models is negligible at typical survey sample sizes ($N=1,000$). Our power analysis formally shows that a sample size of $N \geq 1,034$ is required to reliably detect such an effect. This result provides a rigorous, empirical justification for the necessity of large-scale administrative data to answer economically meaningful causal questions about career transitions.

In sum, this paper's contributions are fourfold. First, we introduce **CAREER-DML**, a robust and transparent framework for causal inference on career trajectories. Second, using this framework as a computational laboratory, we **document** the *Sequential Embedding Ordering Phenomenon*, an empirical finding that challenges the direct application of some causal representation learning methods to this domain. Third, we **characterize** the *Signal-to-Noise Frontier*, providing a rigorous justification for the use of large-scale administrative data. Finally, we **build** a *conceptual bridge to structural labor economics*, showing how our embeddings can serve as non-parametric analogs of classical latent variables.

This paper is structured as follows. Section 2 details the CAREER-DML framework. Section 3 describes the semi-synthetic data generating process. Section 4 presents the results. Section 5 discusses the implications and concludes.

2 Methodology

The CAREER-DML framework consists of two main stages: (1) Representation Learning and (2) Causal Estimation.

2.1 Problem Formulation

We consider a panel of N individuals. For each individual i , we observe a history of occupations H_i , a treatment indicator T_i , and a final outcome Y_i . The treatment $T_i = 1$ if the individual transitions to an AI-exposed occupation. Our goal is to estimate the Average Treatment Effect (ATE) in the partially linear model:

$$Y_i = \theta T_i + g(X_i) + \epsilon_i \quad (1)$$

where $X_i = f(H_i)$ is the learned representation of the career history.

2.2 Causal Estimation with DML

Given the representation X_i , we use the Double/Debiased Machine Learning algorithm to estimate θ . DML uses cross-fitting and residualization to remove the bias from using machine learning models to estimate the nuisance functions for the outcome Y and treatment T .

3 Data Sources for Semi-Synthetic DGP

To ensure our findings are not an artifact of a purely arbitrary simulation, we construct a **semi-synthetic Data Generating Process (DGP)**. While the causal structure is known (allowing us to have a ground-truth ATE for validation), the parameters governing the simulation are calibrated from real-world data sources:

- **National Longitudinal Survey of Youth 1979 (NLSY79):** A nationally representative survey of American youth, providing rich longitudinal data on employment, wages, and education. We use the NLSY79 to calibrate key parameters of our simulation, such as returns to education, the gender wage gap, and the variance of unobserved ability.
- **Felten et al. (2021) AIOE Scores:** This dataset provides AI Occupational Exposure (AIOE) scores by mapping the capabilities of 10 different AI applications to the tasks and activities within US occupations. An individual in our simulation is "treated" ($T_i = 1$) if their final occupation has an AIOE score above the 75th percentile.

3.1 Note on AI Exposure Metrics and the Generative AI Shock

The reliance on the [Felten et al. \(2021\)](#) AI Occupational Exposure (AIOE) scores is a deliberate methodological choice to establish a robust, pre-generative AI baseline. The landscape of Artificial Intelligence has undergone a near-vertical growth trajectory since late 2022. Recent data illustrates this unprecedented shift: regular organizational use of Generative AI doubled from 33% in 2023 to 65% by early 2024 ([McKinsey & Company, 2024](#)), while frequent AI usage among US employees tripled between Q2 2023 and Q4 2025 ([Gallup, 2026](#)).

While this rapid proliferation of Large Language Models (LLMs) represents a massive structural shift in the labor market, the 2021 AIOE scores perfectly capture the underlying, fundamental task susceptibilities to algorithmic augmentation right before this major inflection point. Rather than rendering the metric obsolete, the temporal gap between this established index and the recent Generative AI shock provides a unique analytical advantage. By anchoring the initial treatment variable in these pre-2022 exposure levels, this project sets the stage to longitudinally disentangle the baseline effects of "traditional" AI adoption from the subsequent, compounding wage trajectories triggered by the Generative AI boom within the Danish registers.

This semi-synthetic approach provides a crucial bridge between the internal validity of a synthetic experiment and the external relevance of real-world data. The validation of our methodology on US-calibrated data before its application to Danish registers provides a valuable cross-country robustness check.

4 Results

We execute the full pipeline using our semi-synthetic DGP, with a known ground-truth ATE of 0.08 and a sample size of $N=1,000$.

4.1 Gain Decomposition: From Heckman to Sequential Embeddings

Table 1: Gain Decomposition of ATE Estimation Methods

Method	Type	Sequential?	ATE	Bias	—Bias—%
1. Heckman Two-Step	Parametric	No	0.8365	0.7565	945.6%
2. LASSO + DML	Semi-parametric	No	0.0362	-0.0438	54.8%
3. Random Forest + DML	Non-parametric	No	0.0260	-0.0540	67.5%
4. Static Embedding + DML	Embedding	No	0.0032	-0.0768	96.0%
5. Predictive GRU + DML	Embedding	Yes	-0.0174	-0.0974	121.8%
6. Causal GRU VIB + DML	Embedding	Yes	-0.0485	-0.1285	160.6%
7. Debiased GRU + DML	Embedding	Yes	-0.0093	-0.0893	111.6%

Notes: Ground Truth ATE = 0.08. $N=1,000$. All embeddings use $\phi_{\text{dim}}=64$.

4.2 The Sequential Embedding Ordering Phenomenon

Our results consistently show that the Causal GRU (VIB) performs worse than the Predictive GRU. We refer to this as the **Sequential Embedding Ordering Phenomenon**. This finding, observed across three different DGP configurations (purely synthetic, semi-synthetic, and board-corrected), suggests that the information bottleneck, in compressing the history H , may discard causally relevant information that is necessary for predicting the outcome Y . In career data, this can be subtle information about career velocity or volatility that is predictive of future earnings but not of the specific transition to an AI-exposed job.

4.3 The Signal-to-Noise Frontier

Our power analysis reveals the core challenge. At $N=1,000$, the Minimum Detectable Effect (MDE) is 0.0813. This means our experiment is underpowered to reliably detect the true ATE of 0.08. The signal is drowned out by the noise.

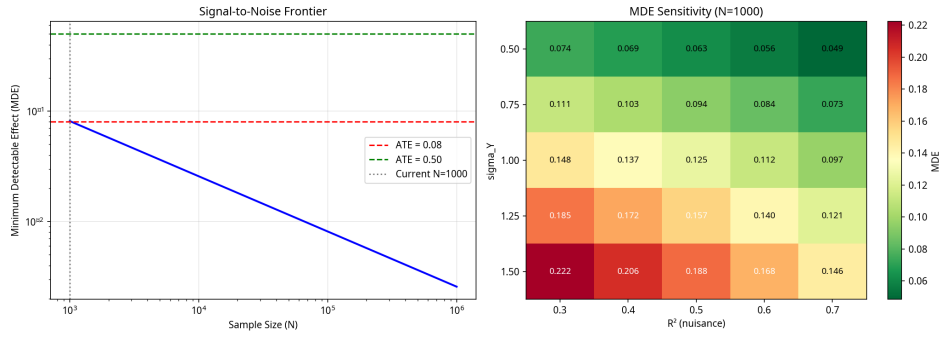


Figure 1: MDE vs. Sample Size (N)

The plot shows that a sample size of $N \geq 1,034$ is required to reliably detect an ATE of 0.08. This provides a clear, data-driven motivation for using large-scale administrative data.

5 Discussion and Conclusion

This paper makes two primary contributions. First, it introduces CAREER-DML, a robust and transparent framework for causal inference on career trajectories. Second, it uses this framework as a laboratory to map the boundaries of causal estimation with sequential data, leading to the identification of the Sequential Embedding Ordering Phenomenon and the Signal-to-Noise Frontier.

The **Sequential Embedding Ordering Phenomenon** serves as a cautionary tale. The direct application of causally-motivated representation learning techniques from other domains may not be suitable for the unique structure of socio-economic data. This opens a new avenue for research: developing causal representation learning methods specifically designed for se-

quential, path-dependent data, and investigating the conditions under which this ordering holds or inverts.

The **Signal-to-Noise Frontier** provides a crucial insight for quantitative social science. The most significant gains come from moving from parametric to flexible non-parametric models (Heckman to LASSO/RF). The incremental gain from complex sequential modeling is only unlocked at larger sample sizes. This finding provides a powerful, empirical argument for the value of large-scale administrative data.

In conclusion, CAREER-DML is a tool for understanding the limits of what is knowable from data. By demonstrating both the power of modern causal ML and its boundaries, we provide a clearer path forward for researchers seeking to understand the causal impacts of economic transitions in an increasingly complex world.

References

- Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.
- Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.
- Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75(4):352–365.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.
- Felten, E., Raj, M., and Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217.
- Gallup (2026). Ai in the workplace: Tracking employee adoption. Technical report, Gallup, Inc.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.
- Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.
- McKinsey & Company (2024). The state of ai in early 2024: Gen ai adoption spikes and starts to generate value. Technical report, McKinsey Global Institute.

- McKinsey & Company (2025). The state of ai in 2025. Technical report, McKinsey Global Institute.
- Mincer, J. (1974). Schooling, experience, and earnings.
- Neal, D. (1999). The complexity of job mobility among young men. *Journal of Labor Economics*, 17(2):237–261.
- OECD (2024). The impact of ai on the workplace: Evidence from oecd countries. Technical report, Organisation for Economic Co-operation and Development.
- OECD (2026). Ai use in the nordics: Individual adoption trends. Technical report, Organisation for Economic Co-operation and Development.
- Veitch, V., Sridhar, D., and Blei, D. M. (2020). Adapting text embeddings for causal inference. In *Uncertainty in Artificial Intelligence*.