The First Generation Conundrum: Model of Wage levels determined by Information Endowment *

Anuva Saran Agarwal
Department of Economics, Duke University

(2023)

[*Working paper]

Abstract

This paper seeks to model a causal link between information endowments obtained via networks and lifetime wage outcomes, conditional on career choice. Information distance, aka social distance in the paper, defined as deviation from existing social/familial networks, determines lifetime wage outcomes through access to information networks. Characterized by information sets, these networks allow individuals to take steps to bridge information gaps. This model is built on the premise that differing information sets distinguish individuals, conditional on a career choice, and for two individuals of the same ability, a difference in their information sets will very likely lead to a difference in the lifetime wages. Wages, aka compensation is based on value-added by an individual. Value added is a function of the individual's potential and actions are determined by information sets where individuals optimize actions to maximise lifetime wages. This paper models differing wage outcomes in career trajectories of individuals who are closer to familial or parental career arenas vs. those who are further away from the career related exposure they have had in their growing years. Under the hypothesis that cost of bridging the deficit in information set is the key distinguishing factor in such wage outcomes, this model concludes that over time the levels of these differences cease to matter however lifetime income may differ due to compounded returns to income as well as information networks.

*I would like to thank Professor Charles Becker, Duke University, for his infinite encouragement and valuable guidance.

1 Introduction

'Social distance' as a determinant of employment and wage outcomes is an observation that all of us (or atleast most of us) would have been confronted with, in our lives and in our interactions with those around us. Think of an economist coming from a family of academics, or better still, from a family of economists. They are born with an 'information endowment' - connections and mentors. Having gained from the experiences of the mentors and interactions in their environment, they would be able to optimize their actions in a career of their choosing should that be similar to the careers of their familial network. To illustrate, anecdotal evidence from multiple conversations with graduate students from top Economics, and Computer Science schools have revealed a common thread - students who come to PhD programs straight from their undergrad seem to have parents or other close familial networks who are in a related field, or have awareness/access to resources that allow one to go to undergraduate institutions where such mentorship exists. The existence of a different engagement (even if it is in the same field) in the midst of their UG and PhD admission could be a signal of a certain cost they have had to pay (payment in time) to accumulate resources - monetary or relevant information. While there are many confounding factors and noise in these decisions, the hypothesis of the existence of such a signal is plausible. Such observations are applicable in various career trajectories - aviators, doctors, lawyers, etc. This paper seeks to model the economic reasoning behind such observed phenomena. Learning is iterative and returns in time are non-linear, for instance the returns to spending time writing content for this paper - low per unit time invested initially, but gradually increasing returns set in and progress became steeper. Psychic cost at each stage in this case would be how far the end output 'feels'. If there are two individuals one of whom believes they can most certainly write the paper vs. another who has doubts, the psychic cost of the second one is higher - the 'feeling of being far away'.

An individual's potential is a function of actions take at each time period T, which depends on the information available, which in turn is a function of distance from access to such information, which has been modelled as social distance. It can be thought of as a measure of struggle – deviation from prior exposure in lifetime and expectation. Information asymmetry induces a certain cost of information acquisition which is a function of 'distance' of the child from parental/familial field of study. How does one get to know, what they don't know? Learning is an iterative process and these are iterations over time, as determined by the number of interactions with people, situations, epiphanies, etc. We model this as degree of exposure. Higher the distance i.e. further the child from previous exposure since childhood, higher is the cost of information acquisition. Additional factors that form a part of this cost are psychic cost, monetary costs, etc. Psychic cost captures the idea that something seems far. Costs of information acquisition are paid in time. An instance of observable indicator of psychic cost, for example, would be longer adjustment time for first generation individuals in any high performing team or students in a university environment.

This paper focuses on the non-monetary costs arising due to differences in 'information endowment' handed by the preceding generations, contemporaneous networks, or both. The question of access to resources (monetary costs for instance) is explored later in the paper. Gains from trade, here trade of information leading to information arbitrage - since it is a public good, it doesn't get 'exhausted' and hence information sharing increases the aggregate information endowment of a group. On the assumption that all agents are capable of converting that information into positive personal growth, group productivity is increasing in information sharing.

The results of this paper point that there is merit in interventions undertaken to facilitate information arbitrage and find ways to reduce the cost of information acquisition (search cost) and psychic cost in the interest of increasing aggregate productivity or value. The applications of this model are far and wide - income distribution implications, the role of peer-to-peer learning in information arbitrage as a means to reducing social distance, as well as providing a rationale for institutional investments to create an enabling environment to facilitate such arbitrage via leadership sensitization and peer sensitization. For instance if a recruiter or a professor comes across a student whose signals and life trajectory seem to be distant from a cookie-cutter model of success, social distance would be an unobservable they must take into account while thinking about such cases.

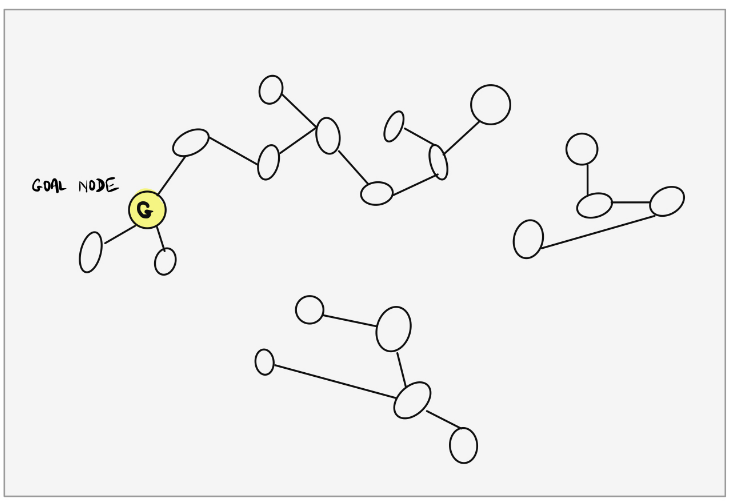
The model built in the following sections build up to conclude that wage differences cease to matter over time. However due to the nature of returns on financial capital (via wage income) and social capital (networks and information set), the expected value of the subsequent generation's income growth is bounded in probability - high deviations (think rags to riches, metaphorically) from this expected income are tail-end events of this probability distribution. Section 2 introduces the model which links social distance to information access via networks. Information endowments are carried over from each time period to the next time period and drive individuals to take decisions in the interest of maximizing their lifetime wages under the assumption that wages reflect the true productive potential of an individual.

2 Model

To simplify analysis the paper divides the spectrum of heterogenous (by career) people into two groups – Type I and Type II. Type I are those whose career or field of study is close to that of their parent(s)' or familial network. And Type II are those who have chosen to pursue careers that are far from what they have seen growing up, and this is the category I am referring to as the 'first generation'. I attempt to expand the definition of first generation to beyond just first generation college students. Type I is within a threshold distance from familial networks and Type II is beyond that threshold. It is this distance that determines information sets based on which individuals take actions to maximize their potential.

This paper seeks to model the higher costs that Type II has to pay relative to the other type. These costs encompass psychic cost c_p and cost of information acquisition c_t with payment in time i.e. $C = \{c_t, c_p\}$. As will be referenced later, these two costs come into the picture when attempting to move between information nodes that do not have any edge between them.

Figure 1 - Information Network



Type 1 = not first-generation individuals; we assume $c_p = c_t = 0$ for Type I since costs can be normalised here. This is the network of nodes connected by edges, to the goal node G.
Type 2 = first gen; characterised by a network of nodes such that no node has an edge between any node of a network containing the goal node G.

Objective function for each time T:

$$\max_{a \in actions} \text{Potential}_T(a)$$

Where action ***a(.)*** is a composite function = ***a(I(d))***; Information set I is a function of social distance d.

The cost of information acquisition is strictly monotonically increasing (since measurement is in time) however psychic cost may or may not be strictly increasing since it may be $= 0$ as well, but is always ≥ 0.

$$\frac{\partial c_t}{\partial T} > 0 \tag{1}$$

$$\frac{\partial c_p}{\partial T} \leq or \geq 0 \tag{2}$$

We allow for the second quantity above to be positive or negative to account for the case when you're so motivated that searching actually fuels you further (i.e. the payment actually feel like inflows). In such a case, c_p which is the feeling of how far the goal seems, is $= 0$ since i no longer feels far from G. c_t is still > 0 and hence, if i is motivated, $c_p = 0$. Hence, $c_p \geq 0$ is a plausible assumption. $\therefore c_p \geq 0, c_t \geq 0$,
$\implies \sum_{costs} C = c_t + c_p = \lambda \geq 0$

Wages are seen to be a measure of potential i.e. the employer is compensating as per the potential of the individual. Hence, without loss of generality, the objective can be redefined as

$$\max_{a \in actions} \text{Wage}_T(a) \text{ for } a \in \text{Action} \tag{3}$$

For the optimizing agent, it is useful to know the actions that will allow them to maximise their wage i.e.

$$\arg \max_{a \in actions} \text{Wage}_T(a) \qquad \forall\ T \tag{4}$$

For lifetime wages, the agent is maximizing $\sum_T \text{Wage}_T(a)$ in which case short term lower wage for a promise of higher future wage can be incorporated – more realistic since wage may not grow linearly as learning curve is a component. We assume that individual i is not present biased i.e. discount factor $\beta = 1$. However, incorporating a discount factor $\beta < 1$ would be a straightforward extension that would not change the results (actions that maximise lifetime wages), here. The objective in that case would be modified to $\sum_T \beta^T \text{Wage}_T(a)$

2.1 Information Space

The information space is a network constituted by information relevant to individual i for compensation commensurate to their best potential. Sources of information are nodes in a network. Learning is an iterative process where iterations can be defined as traversal from one node to another.

Different goal nodes are characterised by different information spaces, which are not necessarily mutually exclusive – there may be a non-empty intersections. This model would apply to each information space, characterized by the goal node. The wage outcomes from the union of all goal nodes would be the set of global maximas of all individual i's potential.

The forthcoming Venn diagrams represent the universe of information.
A = Set of 'Things I don't know'
B = Set of 'Things I know'
$A \cup B$ = Relevant information from the universe of information; refers to information space in Figure 4
$(A \cap B)^c$ = Irrelevant information from the universe of information; outside the scope of this paper
$A \cap B$ = Set of 'Things I know, I don't know'; corresponds to Y axis in Figure 4
$A \cap \neg B$ = Things I don't know, I don't know; corresponds to Z axis in Figure 4
$B \cap \neg A$ = Things I know, I know; corresponds to X axis in Figure 4

Figure 2 - Information Set 2-D Venn Diagram

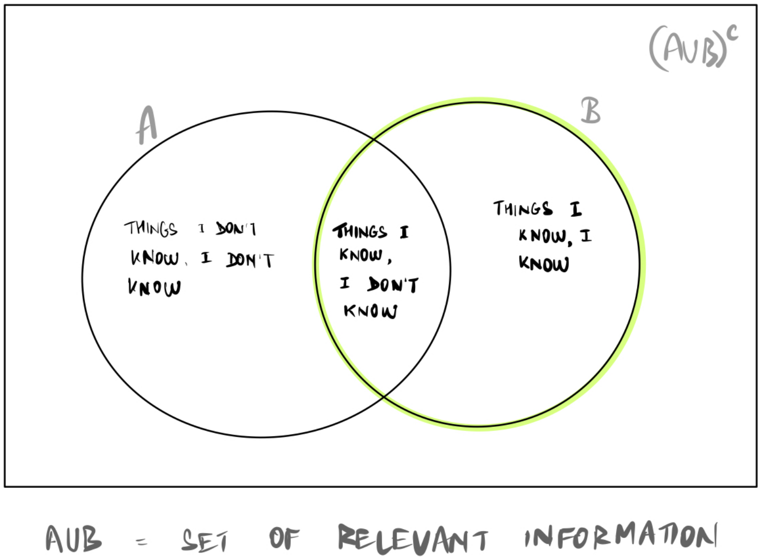


Figure 3 - State of $\tilde{I}$: Information Set A = Set B i.e. $A \cap \neg B = \Phi$

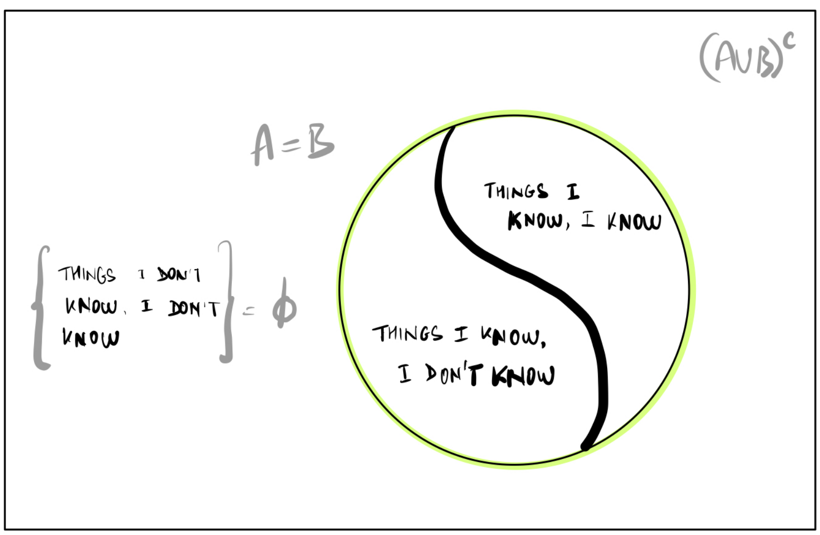
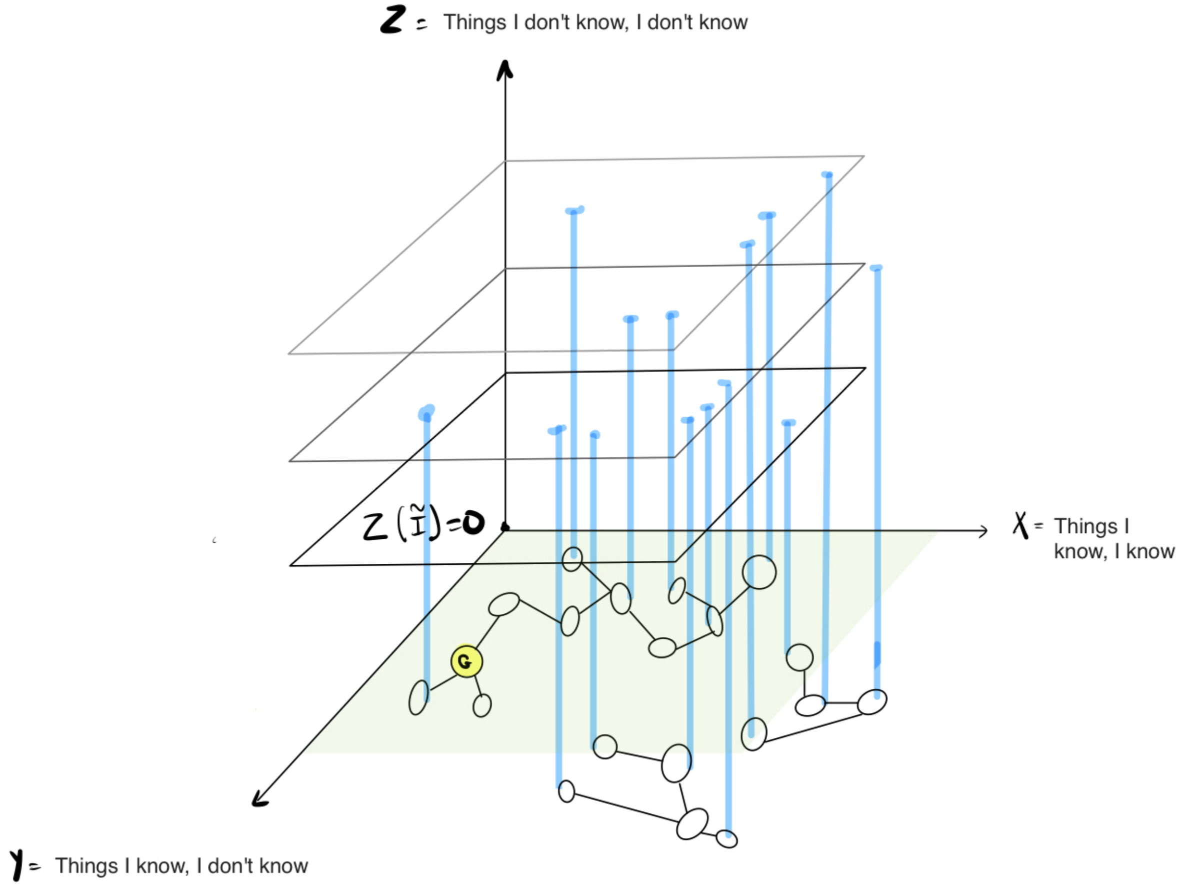


Figure 3 represents a state of $\tilde{I}$ where i knows everything that there is to know, to reach G. We define a continuum of such spaces (as in Figure 2 and 3, both) for each Ⓖ.

Figure 4 - Information space: $A \cup B$ in 3-D.



Each point in the 3-D space in Figure 4 represents I as a function of 'Things I know, I know', 'Things I know, I don't know', and 'Things I don't know, I don't know'. We are always optimizing in the X-Y space since we do not know what level of Z(.) we are at. $Z(\tilde{I})$ corresponds to Set $A =$ Set B i.e. individual i knows everything there is to know, to reach $\tilde{W}$. G is in X-Y space for $\tilde{Z} = 0$ No assumptions on monotonicity or convexity of sets in the X-Y plane. We can think of $Z(.)$ as $\tilde{I} - I$. If $Z > 0 \implies \tilde{I} > I$. i.e. as I increases and come closer to $\tilde{I}, Z = 0$ i.e. we hit the bottom most 'layer' of the X-Y plane.

The basis of this model is the differing information sets between Type I and Type II individuals. The action strategy space for individual i will depend on whether they are in group Type I or Type II – the strategy space will be different. For the purpose of this analysis, WLOG, we collapse all nodes onto one dimension, rank ordered according to the quantity of information acquired. This quantity is measured as the number of nodes traversed to reach a goal node. There could be multiple methods of reaching the goal node but access to that goal node depends on whether or not you have access to that network and this network (information set) has a one-to one mapping with social distance. There exists a kink or discontinuity around the point where Type I would have to 'jump to', to reach the information levels of Type II in order to reach G. This jump is driven by λ (refer to page 3 to recall meaning). $\tilde{I}$ represents everything that one needs to know, to reach G. There may or may not exist and individuals who have access to and this makes the model more generalizable. However, from an observational perspective, the best information that one has about the signal of 'everything that needs to be known' is based on such observed values and hence it is reasonable to assume that $\tilde{I}$ is that which compensated individual i with wage $\tilde{W}$ for a given goal node G. You reach G (goal node) only when you know everything you need to know. Reaching $\tilde{I}$ is a necessary (but not sufficient) condition to reach the goal node since access to resources too plays a role - an extension that has been incorporated in Section 2.5.

To maximise the sum of aggregate wages across all i all individuals must be compensated according to their potential. The nodes can be thought of as points in space representing different volumes of information, ranked as contour plots in the X-Y plane for different values of 'I don't know what I don't know'. Except that they are not level curves but are level 'points'. The vertical distance between different planes where $Z \neq 0$ and the base plane where $Z = 0$ is a monotonically increasing function of λ i.e. as λ falls, individual i moves closer to $Z = 0$. Note that $\lambda = 0$ doesn't guarantee being on the $Z = 0$ plane, it just guarantees being in the information set of Type I which could be at $Z \neq 0$.

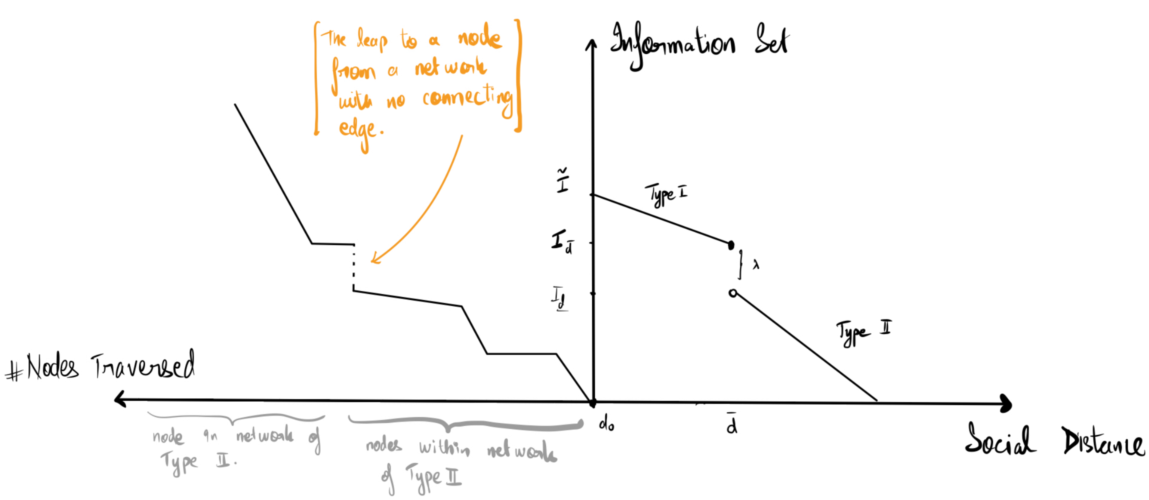
In an ideal world we should work on this model without assumptions of monotonicity since there could be multiple paths to get to the same outcome G with convergence in trajectory but for simplicity, we have ranked information sets by social distance.

2.2 Mapping Social Distance to Information Set and Actions

An action can be defined as steps taken to acquire more information by moving from one node to another via connecting edges (imagine a zoomed in version of the information space). Each node contains information. Nodes that do not have an edge between them, cannot be traversed without incurring an additional cost λ (comprising of c_t and c_p). It is a non-decreasing function since traversing more nodes classifies as more action. However, there is no guarantee that every extra action leads to an increase in I (Figure 5). WLOG we re-design the information space to comprise of nodes that contain unique information and this will allow us to make actions increasing in I. The space of I is restricted to the super-set of nodes that are relevant to all individuals targeting goal G which means that any information is not affecting individuals negatively.

Figure 5 is to be read from right to left: The portion of the horizontal axis to the right and left of the vertical axis represent different quantities - Social distance and Number of nodes traversed, respectively. Note that the gap between the information set of Type I and Type II is given by λ. The number of nodes traversed may or may not increase the information set (horizontal parts of the curve on the LHS quadrant) if the nodes do not provide any new information. However as stated earlier, this space is transformed to one without these horizontal step-like jumps WLOG, but this still allows the curve to be piece-wise linear and continuous.

Figure 5 - Mapping Social Distance to Information Set to Actions



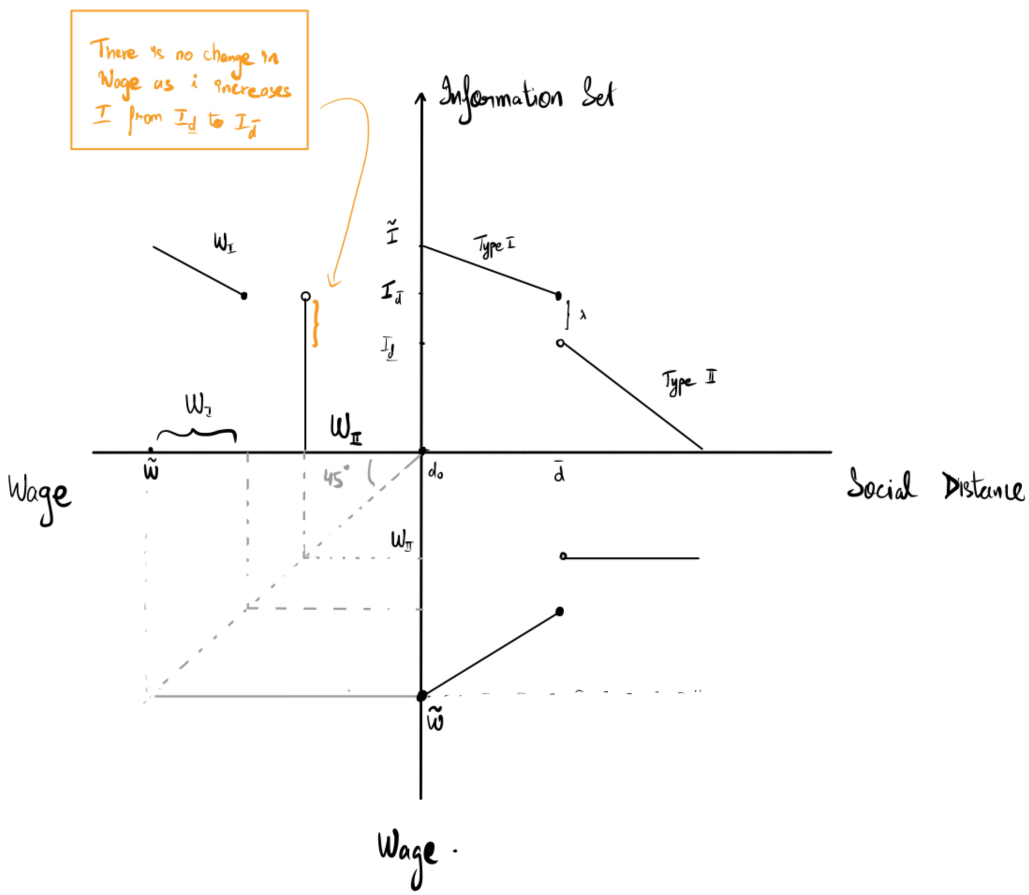
A note on the importance of λ: The jump from Type I to Type II's network involves all sorts of 'jumpy' ways to access the information needed to move to the next network. In context of Figure 4, an i on higher level of Z (further away from $\tilde{I}$) will try to move closer to $Z = 0$ and in the process incur λ. This encompasses getting to 'know what they don't know' and then finding ways to act upon that information in the X-Y plane at the new Z level i.e. as a combination of the 'things that they know' and the 'things they know they don't know'. Whether or not Type I is able to reach Type II's network will depend on search costs too - results from Courant, Paul 1978 paper on housing market can be applied here such that people will search to the point where the foreseeable benefits exceed the search costs incurred. To summarise, the question of whether or not individual i is willing to pay that cost would depend on:

- Time value of money visible in the discount factor in lifetime wage income
- Search costs
- Degree of present bias in agent (prospect theory - Kanheman and Tversky. 1979).

2.3 Mapping Social Distance to Information Set and Wages

Figure 6 is to be read counter-clockwise starting from the top right quadrant. There is a unique $\tilde{W}$ for each $\tilde{I}$ i.e. once someone knows everything they need to know for goal node G, they will not be compensated anything less than $\tilde{W}$. The region marked in orange (top left quadrant) for wage level W_{II} is an important result that states - wage will be the same during the 'search process' before reaching $I_{\bar{d}}$ which is the information set that connects you to a node into the Type I individuals' network, which is necessary to reach G. This mapping carries into quadrant IV at the bottom right, to show differing wage levels as a function of social distance. W_{II} is a single value for social distance $> \bar{d}$ for purposes of visual simplicity. However this can be a decreasing function of social distance too. The key here is the Information (and hence Wage) jump at $\bar{d}$.

Figure 6 - Information Set and Wages as a function of Social Distance



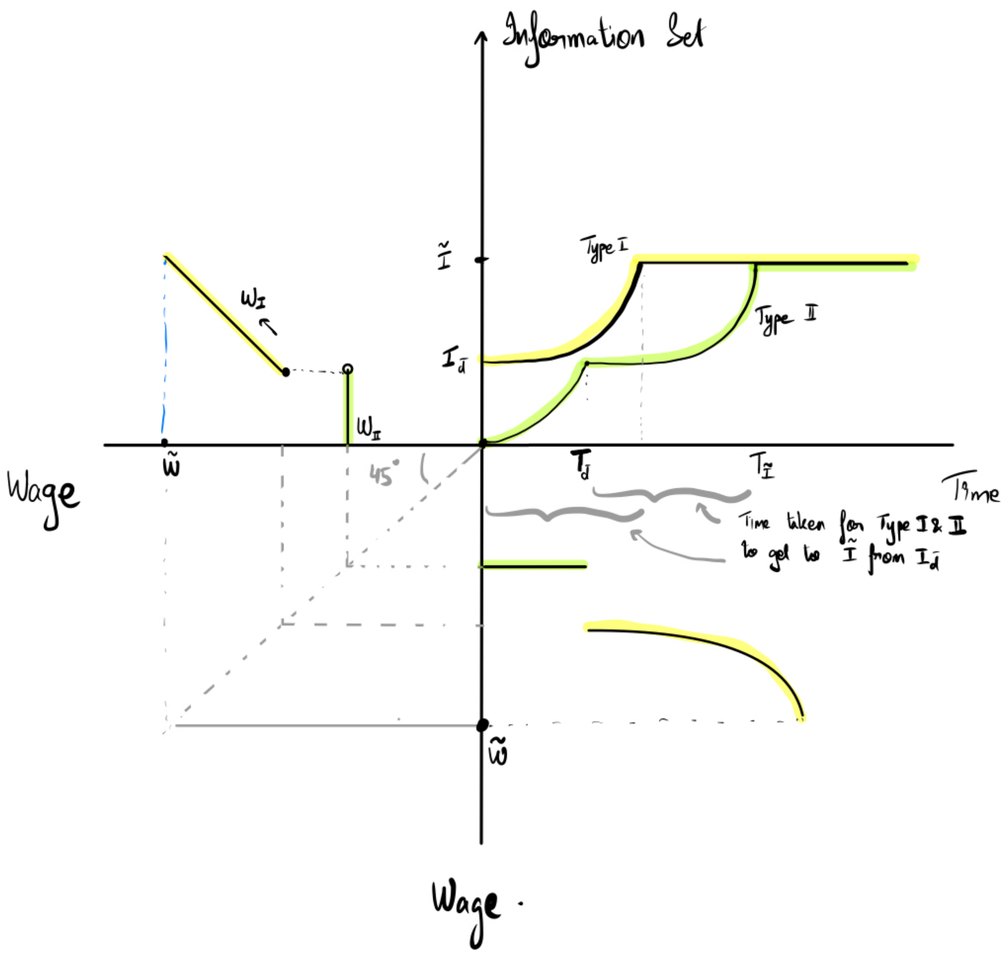
W_I and W_{II} are sets, an interval and a singleton set, respectively. There exists an 'information kink' at the threshold distance (see top right quadrant of Figure 6), where the vertical line represents the additional information that Type II must acquire to reach the 'information levels' of Type I. Each of the quantities in Figure 6 are not time invariant. Any movement along these axes means potential movement with changing time T. Time is modelled in Figure 7 below.

2.4 Mapping Information Set to Time and Wages

Figure 7 is to be read counter-clockwise from the top right quadrant. The graph represents the growth of I across time for Type I and Type II. Key result - wage grows linearly in information set but non-linearly in time. Also note that if there is no change in W at $I > I_{\bar{d}}$ there will be no incentive to be at $\tilde{I}$, and hence wage is increasing in the Time-Wage space after time $T_{\bar{d}}$.

$T_{\bar{d}}$ is the time period at which Type I reaches information level $I_{\bar{d}}$, Notice that at $T_{\bar{d}}$ the information levels of Type I and Type II are different - the yellow curve is above the green curve for Type II. $T_{\tilde{I}}$ is the time period at which Type II reaches information level $\tilde{I}$ and hence reaches wage $\tilde{W}$ (see $\tilde{W}$ labelled on horizontal axis of top left quadrant). This mapping has been carried into the bottom right quadrant to show the time-wage space under these information constraints.

Figure 7 - Information Set over Time as a driver of Wage levels



2.5 Access to Resources and Income Distribution Implications

In an ideal world, once you have everything you need to know, you should be all set. However, differing familial networks may make the 'goal node', here the wage, associated with a specific outcome, seem farther than it actually is (psychic cost). Recall that psychic cost is the feeling of how far a goal node is, and hence psychic costs limit the action space of Type II i.e. for the same action space, Type II incurs a higher cost.

Even if we reach $I(\bar{d})$ there is no guarantee of $c_p = 0$.

In this section we make the above results conditional on access to resources i.e. even if you know everything you need to know i.e. $I = \tilde{I}$ and marginal payment $\lambda = 0$, there is no guarantee of being about to reach W_{II}. This is for Type II individuals in particular. We assume Type I has access to all the resources they need since they are at the point where they can reach goal node G. Even if you believe you can get everything (psychic cost = 0) monetary or other constraints may prevent you from acting upon this information. Access to information is a necessary but not sufficient condition to reach G.
There exist bounds on upward mobility in probabilistic terms, i.e. rightward shifts in the income distribution depends on the probability distribution of access to resources, conditional on information set I.

Proposition 1 - Information **I** and Access to resources are not independent events:

$$P(I|access) \neq P(I) \neq 0$$

$$P(access|I) \neq P(access) \neq 0$$

Information about one will (hopefully) give you an idea about the level of the other variable. Both factors are not independent of each other and are likely positively correlated: more information (recall definition of information that it constitutes relevant information only) would be a function of quality of network and hence a lower λ in moving toward the network of Type I people.

2.5.1 The Role of Scarcity: Why Having Too Little Means so Much

Using insights from the book 'Scarcity' (Shafir and Mullainathan, 2013): the jump from W_I to W_{II} owing to psychic cost, is not a straightforward one. If it consumes all of your bandwidth (mindspace), decision making will be suboptimal, which means that individual i will not be choosing actions which will maximize Wage(a) for $a \in$ Actions. This effect is compounded by the absence/insufficient of access to resources – scarcity of mindspace, sub-optimal decisions in the face of procrastination, high search costs and hence stopping earlier than the average individual i in Type II with a lower search cost. Sub-optimal decisions i.e. actions even with information set $I(\bar{d}$ do not guarantee convergence to information $\tilde{I}$ and hence wage $\tilde{W}$.

Compounded returns on social capital - networks and information access, and on financial capital - wages, would exacerbate these difference and may make $\tilde{W}$ difficult to reach, probabilistically, for Type II. Stories of rags to riches (here 'riches' referring the combination of information set and access to resources) are one-off events and cannot be the norm, statistically speaking.

3 Conclusion and Afterthought

3.1 Extensions for empirical estimation

While not an exhaustive list, the results of this paper may differ from industry to industry.

- Tuition fees – proxy for monetary equivalent of psychic cost – how much would you be willing to pay, to reduce that psychic cost to zero? Once you graduate from Duke, social distance may cease to matter, all other signals equal and assuming there is no heterogeneity on the quality parameters being assessed here.

- Gap between Undergraduate education and Graduate school, controlling for future goals and physical distance between last educational institution and present education institution as an indicator of social distance.

- Intergenerational wealth growth has a probability distribution where there are bounds on how far you can go from the mean income of the preceeding generation.

- Estimate λ by documenting adjustment times for pre-covid/during covid/after covid cohorts. The supportive environment and isolation will capture some effect of this λ. The question is what a good indicator for measuring adjustment times would be.

- What would be an empirical indicator of 'cost' when we say cost of information acquisition, cost of adjustment where payment is in time. Especially relevant in the presence of dynamic continuum of equilibria, each of which are not necessarily independent of each other. By dynamic continuum of equilibria I am referring to maximization of wages in each time period T subject to the constraint that it would maximise lifetime wages.

4 References

1. Spence, M. (1973) Job Market Signaling. Quarterly Journal of Economics, 87, 355- 374.
2. Mullainathan S, Shafir E. Scarcity: Why having too little means so much. New York, NY, US: Times Books/Henry Holt and Co; 2013:289.
3. Coşar, A. Kerem, and Pablo D. Fajgelbaum - "Internal geography, international trade, and regional specialization." American Economic Journal: Microeconomics 8.1 (2016): 24-56
4. Courant, Paul N. "Racial prejudice in a search model of the urban housing market." Journal of Urban Economics 5.3 (1978): 329-345.
5. Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." Econometrica 47.2 (1979): 363-391.