The Potential of Good Selection: Simulating the Impact of Improved Applicant Matching on GDP

By COLIN I.S.G. LEE AND PIERS STEEL*

We estimate the potential GDP impact of increased sophistication during personnel selection and the better use of human capital. In our simulations of the US labor market, selection practices modelled after the standards of selection from 2015 contributed between $1.77 billion (15.2%) and $3.02 trillion (25.84%) to GDP over random selection. A full-scale selection system, using assessments on General Mental Ability and the Big Five personality factors, further added $831 billion (7.12%) to $1.41 trillion (12.08%). Finally, a system using the estimated ideal test battery added $1.10 (9.44%) to $1.86 trillion (15.96%) over the simulation of selection practices from 2015.

* Lee: Amsterdam Business School, University of Amsterdam, Plantage Muidergracht 12, Amsterdam, 1018 TV (e-mail: c.lee@uva.nl). Steel: Haskayne School of Business, University of Calgary, 2500 University Drive NW, T2N 1N4 (e-mail: piers.steel@haskayne.ucalgary.ca). This paper was supported in part by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC). In order to provide transparency, any ties with the assessment industry have been disclosed in the following statement <DISCLOSURE STATEMENT URL>.

The immense benefits of maximizing the value of human capital is among economics' first observations, stretching back to Adam Smith. Specifically, his insight was the productivity benefits brought through the division of labor; as people specialize and increase in expertise, efficiency also increases. However, efficiency grows not only by specialization, but also by the choice of specialization; innate individual differences contribute to performance over and above job training and experience (David Z. Hambrick and Elizabeth J. Meinz, 2011, James J. Heckman and Tim Kautz, 2012). Those blessed with superior dexterity, endurance or intelligence would ideally be placed where these natural talents could be used to full effect. Consequently, our dinners may arrive through the self-interest instead of the benevolence of the butcher, the brewer or the baker, but the quality of that dinner depends partly on an individual's choice of vocation. Even after years of training and experience, variation among innate talents ensures there remains good bakers and butchers as well as bad ones.

Making optimal matches between job candidates and positions is hindered by informational asymmetry, particularly adverse selection as per George A. Akerlof's (1978) "Market for Lemons." Though self-awareness is itself imperfect, job applicants still know more about themselves than their potential employers, or in Andrew Weiss' (1980) terms, "firms have imprecise information concerning the labor endowment of particular worker" (542). This contributes to reduce mobility and frictional unemployment, as employers are less willing to accept the risk of unclear matches (Bruce C. Greenwald, 1986). Consequently, personnel assessment, such as work samples or personality assessment, help mitigate informational asymmetry, potentially boosting economic output. As a further benefit, if done well, with a focus on expected turnover and productivity, assessment minimizes group biases that create adverse impact, where people are discriminated based on non-performance criteria such as race or religion (James L. Outtz, 2010).

Despite their importance, typical selection practices could be far better. A half century ago, Ross Stagner (1958) documented "The Gullibility of Personnel Managers," where employers were typically unable to identify and utilize effective selection practices. Today, this lack of sophistication continues, with a strong preference for "intuition-based hiring," where they rely on subjective, "gut feelings" over more objective, analytic methods. This results in the most popular selection method also being among the least effective: the unstructured job interview (Dirk D. Steiner, 2012). Consequently, considerable improvements can be made in the labor market, with Robert E. Ployhart (2012) arguing that decades of advancement in the selection sciences are not being exploited and there is a huge gap between what has been developed and what is used.

I. Estimating the Potential of Selection Science

In this paper, we estimate the value in the US labor market that is still trapped due to suboptimal selection techniques, drawing on an inter-disciplinary approach as advocated by Janice C Molloy and Jay Bryan Barney (2015). Using a detailed simulation of the US labor market, we examine the extent to which increased sophistication in personnel assessment could improve productivity and, subsequently, GDP.

Estimating the economic impact of selection techniques has a long history, going back to Hubert E. Brogden (1949), who applied cost accounting to personnel selection. Referred to as "utility equations," Brogden calculated the benefits of selection as primarily a multiplicative function of:

1.) the relationship between the selection test and performance, 2.) the increased productivity of an individual performing at one standard deviation above the mean, and 3.) the average standardized test score of those selected from the applicant pool. Thus, as we become better at predicting performance, as the difference between an average and a superior performer becomes more valuable, and as we become pickier or more selective (i.e., increasing the average test score), the worth of our selection system increases. If we cannot predict, if individuals are largely interchangeable, or we must hire everyone who applies, our selection system becomes worthless.

Estimates of the benefits of selection at a national level, though, have been controversial. The U.S. Department of Labor Department requested that the National Research Council investigate the benefits of broader adoption for selection of an extremely well validated test, the General Aptitude Test Battery (GATB). As Richard P. Phelps (1999) documents, a committee of education professors was formed, who "claimed flatly that there are no job selection benefits to testing because the U.S. labor market is a zero-sum game" (38). They argued that when multiple firms are competing for the same employees, there is no net benefit as a good hire for one company leads to a direct reduction in the quality of candidates available for another. Of note, despite "the zero-sum" description, this is notably different from the lump of labor fallacy. The lump of labour refers to the idea that there is limited amount of job positions available, which has been used to argue against the inclusion of women or immigrants to the labor force.

As applied to selection, the zero-sum argument is also fallacious, making sense under five overlapping conditions. First, we would need to assume that those unemployed or underemployed would not be able to add any value to the economy, which would happen only if there was perfect selection. If there are people unemployed who are superior to those holding a job, productivity would improve if these individuals exchanged places. Second, we would need to assume that the attributes to excel in every position in every company are identical. On the other hand, if the requirements for a sales or customer service position are different from the requirements for lawyers or surgeons, there is opportunity for improved matching. Third, we need to assume that all jobs are equally valuable. Increases in productivity can arise due to "allocative efficiency," where we channel those with most skill where they are most needed. For example, Christopher Langan, whose IQ is among the highest recorded, had a long stint as a bouncer on Long Island. Though he may have been an outstanding bouncer, arguably, his cognitive talents may have been better utilized elsewhere. Fourth, the zero-sum argument adopts a form of "The Just World

Hypothesis" (M. J. Lerner, 1980), where there is no racism or sexism or bias against any group based on non-performance related characteristics. If we acknowledge that bias exists, then good selection practices can partner in minimizing it while improving overall productivity. Fifth, the zero-sum argument could be maintained under the condition that we are all inter-changeable, what might be described as "The Clone Army Hypothesis."

Given the implausibility of any one of these five conditions, there has been an early attempt at providing a national estimate of the benefits of selection . Though somewhat rudimentary, based on using just four occupational groups (i.e., professional, skilled, clerical and unskilled) with a simplified selection model based on a single general ability factor, they applied Brogden's (1949) utility equations to the 1980 U.S. economy. Among several scenarios, they estimated an increase to the U.S. GDP of 15.8%, or approximately $453 billion in 2017 dollars, compared to random selection.

Over the ensuing decades, a series of selection and simulation advances have occurred. In particular, the development of *synthetic validity* has matured, where *validated* selection systems can be *synthesized* from predetermined correlations with job components almost instantly and at negligible cost (Piers Steel et al., 2006, Piers Steel et al., 2010). Previously, the development of full-scale selection system was prohibitively complicated and costly that it was primarily the purview of large institutions with jobs that held positions for hundreds or more. Also, with today's digitalization of applicant data (Eric M. Dunleavy et al., 2008, Paul A. Gilster et al., 2001) and the use of digital footprints in assessment (e.g., Wu Youyou et al., 2015), the unit cost of administering selection tests is approaching zero. This has substantially lowered the threshold to algorithm-, equation-, or automation-based recruitment and selection (Corissa Leong, 2018, Dianna L. Stone et al., 2015). Given that the benefits of mass improvements in selection are now readily realizable, the importance of verifying the 1980 figure is timely. If the 15.8% improvement still holds, that translates into over three trillion dollars in added productivity.

II. Simulation Components

The elements required to simulate the effect of selection on an economy are A.) applicants, B.) vacancies, C.) a method to match these applicants and vacancies, and D.) a technique that allows productivity to be estimated for the selected applicants and, successively, the impact on the economy's Gross Domestic Product (GDP). Before showing how these elements function in unison in a simulation of the U.S. labor market in 2015, we will explain how each separate element was generated. At every step, we focused on the U.S. economy in the year 2015 specifically, by using data from that year for aspects that unmistakably vary over time.

A. The Labor Force

According to the U.S. Bureau of Labor Statistics, there were 153,763,688 individuals in the workforce in 2015. Of these, a little over 5% was unemployed (Bureau of Labor Statistics, 2016). We recreated this workforce in a Monte-Carlo simulation including, for each worker, the key characteristics that are traditionally required to estimate the individual's performance in an occupation.

In applicant selection, intelligence and personality are generally recognized as the most important predictors across occupations (e.g., In-Sue Oh et al., 2015, Frank L. Schmidt and John E. Hunter, 1992, 2004, 1998). We gave our population scores on these characteristics by providing each individual a score such that the population has the same mean, standard deviation, and overall distribution, as was found in prior meta-analytic studies on General Mental Ability (GMA; Ute R. Hülsheger et al., 2007) and the Big Five personality measures (Timothy A. Judge et al., 2013, John P. Meriac et al., 2008), for intelligence and personality respectively. To ensure that the combinations of values reflect those found in the workforce, we spawn the distribution of the values across the workforce from a matrix containing the mean, corrected, correlations (i.e., ρ) and the corresponding standard deviations (i.e., sd_ρ) between the characteristics. This Monte-Carlo process (similar to Patrick D. Converse and Frederick L. Oswald, 2014) results in a sample with correlations and score distributions on GMA and personality, that are effectively indistinguishable from those observed in the prior studies that the selected meta-analyses draw from.

B. Occupations and Jobs

Data on the mean salary per occupation (used to establish productivity, as will be explained later) and number of available jobs per occupation for the year 2015 were retrieved from the U.S. Bureau of Labor Statistics (Bureau of Labor Statistics, 2016). The data on the characteristics of the occupations in the U.S. labor market were obtained from the Occupational Network Online (O*NET), a platform created and managed by the U.S. Department of Labor (for a summary see Norman G. Peterson et al., 2001).

O*NET provides numerous data tables with detailed descriptions of occupations. We used two sources of job characteristics data in particular. The main source was the O*NET listing of General Work Activities (GWAs). The use of GWAs for synthetic matching was recommended by Rodney A. McCloy et al. (2010). We simplified the process by reducing the 42 O*NET GWAs into the 3 well-established job components of "Data," "People", and "Things" – an approach supported by prior factor analytic work (Sidney A. Fine and Steven F. Cronshaw, 1999, Shanan G. Gibson et al., 2007, Theresa M. Glomb et al., 2004). "Data" focuses on problem solving and dealing with information (e.g., "thinking creatively"), "People" focuses on interpersonal relations (e.g., "training and teaching others"), and "Things" involves physical manipulation (e.g., "handling and moving objects").

Expanding on Data, People, and Things, Rustin D. Meyer et al. (2009) and Rustin D. Meyer et al. (2014) developed the additional components of "Constraints" and "Consequences." The constraints component is defined as "the amount of behavioral/decisional restriction placed on an employee or, conversely, as the amount of autonomy or latitude an employee experiences" (Rustin D. Meyer, Reeshad S. Dalal and Silvia Bonaccio, 2009). Consequences is defined as "the presence of contingencies between one's decisions or behaviors and the outcomes accruing to oneself, other employees, the organization as a whole, and/or external stakeholders" (Rustin D. Meyer, Reeshad S. Dalal and Silvia Bonaccio, 2009). Most items for these two components came from a second source, the O*NET work context table.

C. The Selection System

To simulate the evaluation of individuals in applicant selection and to estimate how individuals in the labor force would perform within specific occupations, we built a *Synthetic Validity* selection system. Synthetic Validity determines the weights of the performance indicators for an occupation by calculating how the weights vary (i.e., are moderated) by work activities and context (Jeff W Johnson et al., 2010, Piers Steel, Jeff W. Johnson, P. Richard Jeanneret, Charles A. Scherbaum, Calvin C. Hoffman and Jeff Foster, 2010). Essentially, by examining the task or conditions that comprise a job, we can estimate who would be best to fill the position (i.e., "From the Work One Knows The Worker"). As a rudimentary example, general mental ability is more important for performance in complex jobs that require considerable problem solving (i.e., higher in the Data component). Once we know the magnitude of that effect and the general mental ability score of applicants, we can get a first estimate of their performance. This can then be extended by including other indicators and taking into consideration the overlap between these indicators (i.e., intercorrelations).

In the simulation, each occupation has been scored on the extent to which the job components Data, People, Things, Constraints, and Consequences are required or needed to perform the occupation. As described, these scores were retrieved from the O*NET database. Knowing the activities performed in each occupation, we only needed to determine how the validity coefficients for each predictor (GMA and each of the Big Five personality components) correlated with those activities (reduced to the job components Data, People, Things, Constraints, and Consequences), to establish occupation-specific regression equations.

Validity Estimates —To determine how the validity coefficients of the predictors correlated with the job components, we relied on 54 *subject matter experts*, personnel selection scholars and practitioners from the field of Industrial and Organizational Psychology, and an elaborate assessment environment. Prior studies using clinical or expert predictions within applied psychology and medicine, have shown that subject matter expert estimates are typically accurate and robust (William M. Grove et al., 2000). In a comparison of subject matter expert estimates of validity coefficients in applicant selection with coefficients derived through ordinary least squares regression, O'Neill and Steel (2018) reveal that we would need validation and job analysis data from at least 185 jobs to provide an estimate as precise as that by 10 experts. That is, the uncertainty of empirical estimations due to sampling error is evenly matched by the subject matter experts' imprecision (i.e., both forms of estimation create correlations that are within .04 of the true 50% of the time). Notably, we relied on 51 subject matter experts for our estimates in the present study. The initial sample consisted of 54 experts, but 3 were excluded because they failed to follow the instructions and provided multiple estimates outside the specified ranges. Also, the 185 jobs represent level 2 of a multi-level model, with level 1 being the selection system for the individual jobs, each of which, if empirically derived, would essentially be a separate study composed of data from hundreds of employees. Consequently, the data derived here is equivalent to some of largest studies ever attempted in the personnel selection sciences (Piers Steel et al., 2006).

In order to ensure high data quality, we followed the procedure for online subject matter expert estimation of Synthetic Validity job component correlates, as outlined in detail by Rodney A. McCloy, Dan J. Putka and Robert E. Gibby (2010). We then refined the approach by restricting the experts' estimates to ranges that were viable statistically, considering the known meta-analytic distributions of the validity coefficients from the ρ and sd_{ρ} values reported by Ute R. Hülsheger, Günter N. W. Maier and Thorsten Stumpp (2007), Timothy A. Judge, Jessica B. Rodell, Ryan L. Klinger, Lauren S. Simon and Eean R. Crawford (2013), and John P. Meriac, Brian J. Hoffman, David J. Woehr and Matthew S. Fleisher (2008). With this we arrived at informed, conditional estimates using *Meta-Analytic Assisted Subjective Estimation*. For a full explanation of the procedure, see Appendix A. *Job Specific Assessment Equations* —In order to derive job-specific equations that would be used in the assessment of an applicant for a specific vacancy, we used the latest approach to Synthetic Validity as described by Piers Steel, Allen I. Huffcutt and John Kammeyer-Mueller (2006) and demonstrated by Piers Steel and John Kammeyer-Mueller (2009). When incorporating subject matter expert estimates instead of empirical weights, the approach consists of roughly three stages.

First, the means of the scaled subject matter expert estimates for the second-order correlations (i.e., correlations between predictor's validity coefficients and the job components) were used in six separate OLS regression (one for each predictor), to obtain a regression equation that allows the validity coefficients of each predictor to be estimated from an occupation's job components. Second, these regression equations were used to estimate the job-specific validity coefficients for each of the 874 O*NET occupations listed with full data on O*NET – this excludes 20 O*NET occupations that are specific to the military, 124 occupations that serve as general or rest categories (e.g., "Engineers, All Other"), and 82 occupations for which the Bureau of Labor Statistics does not keep employment data (e.g., "Baristas", "Endoscopy Technicians", "Chief Sustainability Officers"). In the third and final step, these job-specific validity coefficients were entered into a correlation matrix containing the correlations between the predictors (i.e., as in Table 1 in Piers Steel, Allen I. Huffcutt and John Kammeyer-Mueller, 2006) and OLS was applied with job performance (i.e., the validity coefficients) as the outcome variable. This provided a unique, synthetic, regression equation for each occupation that allows job performance to be predicted from GMA and the Big Five personality factors.

D. Estimating Impact

We seek to estimate the potential of selection, a scenario where geographic mobility and educational opportunities are not substantive limiting factors. When the simulation runs, the people in the labor force apply for jobs and go through a selection process, where they are tested on their intelligence and personality. Using the job specific synthetic equation, this provides an overall job performance score for the individual on the specific occupation. Since this is our best estimate of the individual's performance, we take this value as the objective score (i.e., supposed "true" performance score) to which we add pseudo-random values with a Gaussian distribution (i.e., "random noise"), with a weight that replicates the variance explained (i.e., R^2) of the regression equation, to reach a predicted performance score of the individual in an application process.

The predicted performance scores are used to allocate people to jobs. After all candidates have been tested, the candidate with the highest score is offered the position. Consistent with the high weight given to salary in job choice (e.g., Derek S. Chapman et al., 2005), individuals respond to incentives and take the job with the highest pay.

Baseline selection system —To simulate the selection proficiency and productivity levels found in the 2015 labor market (i.e., $\Delta\boldsymbol{U}_{\boldsymbol{baseline}}$), we first recreated the *natural selection effects*. Here, this term refers to selection effects that occur before people apply to jobs, which includes everything from parental influence to educational choice. As per several theories of occupational choice, job applicant pools are not random draws from the population but reflect some inherent matching (Seung-Ming Leung, 2008). For example, in order to become an engineer, one needs to graduate from an engineering school, which serves as an indirect filter for the selection process (Christopher M Berry et al., 2006). After having recreated naturally occurring selection, we included selection effects of contemporary selection practices.

Natural selection was simulated by making the average ratio between the Standard Deviation (SD) of GMA within the applicant pools and SD of GMA across the population equal to the ratio observed within applicant pool statistics. Such ratios are known as *SD ratios* (Deniz S. Ones and Chockalingam Viswesvaran, 2003, Paul R. Sackett and Daniel J. Ostgaard, 1994), reflecting between and within group variance. The SD ratio of GMA indicates the possible extent of natural selection, as it reduces when people choose or are sorted into occupations for which they are comparatively well-suited. For example, random selection would indicate that within group variance is the same as between group variance, creating an SD ratio of one.

The SD ratio for the contemporary labor market was established using applicant pool statistics from a widely-used GMA assessment test, the *Wonderlic Personnel Test* (Eldon F. Wonderlic, 2007). The ratio showed that, on average, the SD of applicant pools was 11.8% lower than the SD across the population.

This SD ratio was calibrated in the simulation by increasing the variance accounted for in performance until the applicant pools in our simulation had a similar SD ratio to the ratio observed in the contemporary labor market. This indicated that the variance accounted for by natural selection is at most .049. This is an upper limit, since the approach builds on the conservative assumption regarding the Gravitational Hypothesis (i.e., the attraction, selection, attrition model; Steffanie L. Wilk et al., 1995), that all selection is to the benefit of performance. In practice, as some of the homogeneity in applicant pools is due to non-performance related characteristics (e.g., Christopher D. Nye et al., 2017), there is some natural selection that does not contribute to performance. For instance, vocational counseling often relies on self-rated abilities to help individuals make career choices. Considering the limited relationship between self-rated abilities and actual abilities, this can increase homogeneity in the applicant pool, without improving performance. In this respect, the simulation is conservative as it overestimates the productivity increase that the baseline selection practices provide over random selection and underestimates the potential gains remaining.

After having simulated natural selection, which would include educational requirements, we considered the effects of selection practices used by companies. Companies commonly only use unstructured interviews in the selection process (Dirk D. Steiner, 2012). Therefore, to replicate the variance accounted for by the selection effects in the application process, we increased the variance accounted for by natural selection with the variance accounted for by unstructured interviews. We based the variance of unstructured interviews on a meta-analysis that included 39 studies and had a combined sample size of 9,330 (i.e., .23; Michael A. McDaniel et al., 1994). We employed the validity coefficient without the correction for range restriction (R^2 = .0529), since natural selection already accounted for range restriction effects. The combined variance accounted for by natural selection and selection by the hiring firm at baseline was thus .1019 on average, across occupations ($R^2_{\text{selection}} = R^2_{\text{natural selection}} + R^2_{\text{firm selection}} = .049 + .0529 = .1019$).

Utility analysis —In order to compute the percentage change in productivity ($\Delta\%\boldsymbol{U}$) when contemporary selection is replaced by another form of selection, we used a simplified version of the Brogden utility model (Hubert E. Brogden, 1949, 1946). First, for each simulation of a selection system, the total dollar value change over random selection ($\Delta\boldsymbol{U}$) was calculated by taking the sum total change for every occupation:

(1) $$\Delta \boldsymbol{U}_{\boldsymbol{new}} = \sum_i \boldsymbol{N}_{\boldsymbol{i}} * \boldsymbol{SD}_{\boldsymbol{Y}_{\boldsymbol{i}}} * \overline{\boldsymbol{u}}_{\boldsymbol{X}_{\boldsymbol{i}}}$$

Where for each occupation $\boldsymbol{i}$, $\boldsymbol{N}_{\boldsymbol{i}}$ denotes the number of individuals selected, $\boldsymbol{SD}_{\boldsymbol{Y}_{\boldsymbol{i}}}$ is the standard deviation in job performance in monetary units ($\boldsymbol{Y}$), and $\overline{\boldsymbol{u}}_{\boldsymbol{X}_{\boldsymbol{i}}}$ is the mean estimated job performance score for those selected, where job performance was a z-score prediction from the Synthetic Validity selection system. For $\boldsymbol{SD}_{\boldsymbol{Y}}$ we used both the conservative cut-off of 40% of mean salary (which as Schmidt and Hunter report is "a lower bound value; actual values are typically considerably higher" p. 263) and the more liberal range of 40-70% (John E. Hunter and Frank L. Schmidt, 1982a), where the value increases proportionally with the O*NET job zone of the occupations (John E. Hunter et al., 1990).

Being a simulation and having a purview that goes beyond the typically obtainable components, this approach deviates from the original Brogden model in that it uses the objective performance scores of those selected (i.e., $\overline{\boldsymbol{u}}_{\boldsymbol{X}_{\boldsymbol{i}}}$) to establish productivity, instead of inferring the productivity from mean predicted performance in the applicant pool. This approach enables us to model the distributive effects of selection more precisely, since the utility for the occupation is dependent only upon those selected and not on those in the applicant pool that were not selected.

We established the percentage change in productivity provided by each selection system by subtracting the value added by the contemporary system ($\Delta\boldsymbol{U}_{\boldsymbol{baseline}}$) from the value added by the new system ($\Delta\boldsymbol{U}_{\boldsymbol{comparison}}$) and dividing this by the overall utility in the current system, as reflected by the sum total in salaries ($\sum_i \boldsymbol{N}_{\boldsymbol{i}} * \overline{\boldsymbol{E}}_{\boldsymbol{i}}$):

(2) $$\Delta\%\boldsymbol{U}_{\frac{\boldsymbol{comparison}}{\mathbf{baseline}}} = \frac{\Delta\boldsymbol{U}_{\boldsymbol{comparison}} - \Delta\boldsymbol{U}_{\boldsymbol{baseline}}}{\sum_i \boldsymbol{N}_{\boldsymbol{i}} * \overline{\boldsymbol{E}}_{\boldsymbol{i}}}$$

Where, for every occupation $\boldsymbol{i}$, $\overline{\boldsymbol{E}}_{\boldsymbol{i}}$ is the average income for 2015.

III. Improvements in Selection: Simulating the Impact on GDP

Figure 1 shows the results of the simulations under the conservative assumption of the effect of job performance on productivity (i.e., SD_Y = 40%) and the assumption that the effect of performance on productivity varies with job complexity (i.e., $40\% \leq SD_Y \leq 70\%$). Both trend lines show diminishing marginal returns to the economy of investments in applicant selection (i.e., test battery accuracy). They indicate that investments in applicant selection do lead to a substantial net performance and productivity gain to the economy, but the performance gains diminish as adoption increases.

[Insert Figures 1 Here]

A. Improvements in Selection

As can be seen in Figure 1, we first established the value added of the baseline system compared to a system in which the selection of people to occupations is completely random. The system modelled after contemporary selection practices provided an improvement of 15.2% over random selection, when using a conservative estimate of the effect of job performance on productivity (i.e., SD_Y = 40%). Since labor accounts for approximately 65% of GDP in the US (Douglas Gollin, 2002, Alexei Izyumov and John Vahaly, 2015), we can convert this to a monetary value by multiplying the productivity increase by 65% of GDP in 2015 (i.e., $17.947 trillion; International Monetary Fund, 2016). Thus, the productivity increase from selection practices in 2015 (including self-selection) translates to an estimated $1.77 trillion rise in GDP (i.e., .65*17.947*.152 ≈ 1.77). When assuming that the impact of a one deviation increase in performance leads to an increase in productivity of up to 70% for complex occupations, these estimates rise to 25.84% and $3.02 trillion.

Subsequently, we created a system in which every company assesses its applicants using GMA and the Big Five personality test. In practice, the underlying algorithms could be hidden and a company would simply be informed of the desirability to hire an applicant. In this simulation, productivity was 7.12% higher than in the simulation of the contemporary labor market, under conservative estimates of SD_Y. In terms of a GDP increase this translates to $831 billion (i.e., $0.831 trillion). When SD_Y varied by job complexity, the simulation indicates that a further rise of

12.08% over the 2015 baseline can be expected, which means that selection would by these attributes would have added $1.41 trillion to the US economy.

The consistent use of GMA and Big Five by companies seems to be the most plausible scenario (e.g., Orlando Behling, 1998, Lewis R. Goldberg, 1999). However, one can question whether companies would choose to base their hiring decisions on GMA and the Big Five only. Therefore, we simulated more extensive use of applicant testing by increasing the variance explained by selection in our simulation up to an estimated maximum possible variance that could be explained by any combination of assessment tests. To find this maximum, we asked our panel of subject matter experts to provide their estimates of the variance explained of the "ideal test-battery": "… How much variance in performance do you think we would be able to predict if every psychometric improvement (i.e., the best measures, of the best quality) would be implemented?"

The mean estimated variance (i.e., R^2) of the hypothetical ideal test battery was 50.32% (N = 54 subject matter experts). To simulate the use of such assessments, we increased the R^2 of the test batteries based on GMA and the Big Five proportionally per occupation, from the original vacancy-weighted R^2 average of .3456 to the vacancy-weighted average representing the ideal (i.e., R^2 = .5032). This increased the conservative estimate vis-à-vis the contemporary system to 9.44% and a GDP increase of $1.10 trillion. When SD_Y was contingent upon job complexity, the increase was very similar to the original estimate by John E. Hunter and Frank L. Schmidt (1982a), 15.96% and $1.86 trillion.

B. Improvements in Allocation

In our initial simulations, we assumed that vacancies with the highest salary get first choice in the labor market. This selection mechanism should resemble reality to a fairly high degree, as salary is an important determinant of job choice (e.g., Derek S. Chapman, Krista L. Uggerslev, Sarah A. Carroll, Kelly A. Piasentin and David A. Jones, 2005), but it does not provide optimal allocative efficiency. Some individuals might be comparatively well suited for a lower paid job such that they would contribute more to the economy if they would take the lower paid position.

To test the extent to which allocation could improve productivity, we created a system in which individuals are optimally allocated to vacancies through a stable allocation algorithm (David Gale and Lloyd S. Shapley, 1962, Alvin E. Roth, 1984). The algorithm consisted of four stages. First, companies determined the productivity of each candidate for each occupation and offered the available positions to the best candidates. Second, applicants retained their most attractive offer (but deferred from accepting) and rejected the offers for jobs in which they would be less productive. Third, the companies selected the next-best candidates for jobs on which the prior offer was rejected. Finally, the market went through steps two and three until all of the jobs were filled.

We first tested the increase in productivity that optimal allocation provided along with selection on GMA and the Big Five personality factors over the 2015 baseline, under the conservative assumption of the effect of job performance on productivity (i.e., SD_Y = 40%). This revealed an increase 8.02% or $936 billion (i.e., $0.936 trillion) over the 2015 baseline. When SD_Y varied by job complexity, the increase over the 2015 baseline was 13.44% and $1.57 trillion. Compared to the system in which salary determined the order of allocation, ceteris paribus, the increase was 0.90% and $105 billion (i.e., $0.105 trillion) under the conservative assumption regarding SD_Y and 1.35% and $158 billion (i.e., $0.158 trillion) otherwise.

Next, we tested the increase that would be provided by optimal allocation in the simulation using the ideal test battery for selection. This showed a 10.6% or $1.24 trillion increase over the 2015 baseline under the conservative assumption regarding the relationship between job performance and productivity. When the relationship between job performance and productivity varied by job complexity, the increase was 17.64% and $2.06 trillion. This translates to an increase of respectively 1.16% or $136 billion (i.e., $0.136 trillion) and 1.68% or $196 billion (i.e., $0.196 trillion) over the same selection system with the market (i.e., salary) determining allocation.

C. Returns to Adoption

To provide insight into the extent to which the benefits of applicant selection decrease with adoption, we compared the progression of average productivity in jobs relying on contemporary selection practices versus the average productivity for jobs relying on GMA and the Big Five for applicant selection. We chose to compare to the latter because this is the more realistic projection of the future use of applicant selection in the labor market.

[Insert Figure 2 Here]

The comparison shows that early adopters indeed have substantially higher gains. When the rate of adoption is 10% or lower, productivity is over 32.9% higher in occupations relying on the GMA and Big Five test battery for selection in comparison to the contemporary productivity rates (i.e., relying on a selection system that accounts for .1019 of the variance in productivity on average) and 30.1% higher compared to non-adopters (i.e., occupations relying on a system with an average efficacy similar to that of contemporary selection practices). As adoption increases, there are diminishing marginal returns to adoption, ranging from 39.4% at 2% adoption to 7.1% at 100% adoption in comparison to contemporary productivity rates. The “penalty” of non-adoption ranges from -.005% at 2% to -6% and below after 50% adoption. Of note, when adoption is low, the trend for adopters starts diverging and when adoption is high, the trend for the non-adopters ranges because of the lower vacancy counts.

IV. DISCUSSION

This study shows the potential of employment selection in enhancing national productivity by providing several simulations of the US labor market. We show how productivity and GDP for the US would have improved in 2015, had every company conducted psychometric testing in applicant selection.

Under the most feasible conditions, where companies would test on GMA and the Big Five personality constructs and individuals choose their occupation based on salary, productivity improved by 7.12% to 12.08% and GDP increased by $0.831 to $1.41 trillion. When every company would use the state of the art in applicant testing (i.e., the “ideal” test battery), the productivity improved by 9.44% to 15.96% and GDP increased by $1.1 trillion to $1.86 trillion. Compared to changes in applicant testing, using optimal (i.e., "stable"; David Gale and Lloyd S. Shapley, 1962, Alvin E. Roth, 1984) allocation methods to assign individuals to occupations had a small effect on productivity and GDP, with an improvement of 8.02% ($0.94 trillion) to 13.44% ($1.57 trillion) when testing on GDP and the Big Five and 10.6% ($1.24 trillion) to 17.64% ($2.06 trillion) when using the estimated ideal test battery.

Given the magnitude of these estimates, it is worth reconsidering their verisimilitude. We did not model job mobility, matching job applicants to jobs without regards to geography. The degree that this is a concern is ameliorated when simulating at an occupational level, as a job applicant merely needs to matched to a position within the region defined by their mobility, not a specific firm or organization. The larger the region, such as New York’s tri-state, the more it will reflect occupations representative of the nation as a whole. Modeling representative national sub regions and then summating would be equifinal to our simulation.

In addition, we did not directly model job training or educational requirements, except for its role in natural selection (e.g., Christopher M Berry, Melissa L Gruys and Paul R Sackett, 2006). While education and ability are correlated, they are imperfectly so. Consequently, this simulation reflects a country's educational meritocracy, that is its ability to maximize its human capital in terms of both selection and the underlying education simultaneously. This is sporadically an issue. O*NET sorts occupations into five job zones, each reflecting a degree of preparation or training. Level 1 and 2 require no more than a high school diploma or GED (General Equivalency Diploma) and Level 3 requires another year or two of training. Together, they represent two-thirds of occupations, and given that the U.S. Census Bureau indicates approximately 90% of American's have a high school degree or equivalent, educational requirements pose few obstacles here. The remaining third of occupations represent level 4 and 5 jobs, which require advanced degrees that can be expensive or difficult to obtain. To fully realize the benefits of selection, a country would need to enable those with only a high school degree but selected for occupations in job zone level 4 or 5 to obtain the requisite education. While as a society we collectively try to minimize these instances through vocational counselling and scholarships, as per the case of Christopher Langan, there will be lost opportunities. While our estimates still represent the amount of human capital untapped in the labor market, it may of interest to partial out the portion that would remain unrealized due to a lack of educational meritocracy even with ideal selection practices.

On the other hand, it some ways these are still conservative estimates, based on the maximal impact of natural selection. If, for example, natural selection was midway between random selection and its present maximal, the benefits of improved selections practices would increase by approximately by another 500 billion annually.

Across all simulations, the results reveal that improved selection practices provided a substantial improvement to GDP. Nonetheless, the returns to adoption do depend on the overall rate of adoption in the market. When adoption rates are low, returns on the use of applicant testing on productivity can be as high as 39.9% compared to occupations not adopting the improved practices. At near-full adoption, the benefit of adoption versus non-adoption was estimated at 11.6%.

Consistent with simulation methodology, not all potentially relevant parameters could be included. We did not, for instance, model indirect effects of improved selection practices. Good selection reduces frictional unemployment, increasing the percentage of time in gainful work. It also results in more effective vocationally counselling or self-selection, where people pursue positions not only based on personal interests, but on realistic likelihoods of employment. This would reduce false or errant career paths, maximizing the benefits of training and education. We leave these as questions for future research, though notably both would further increase the estimated benefits of improved selection. Nevertheless, our simulation estimates already provide a good and well-substantiated indication of the impact and potential of large scale adoption of more sophisticated approaches to applicant selection. Given we are well into an era of interconnected and interoperable applicant data, these estimates are particularly relevant as the underlying selection practices we modelled could readily become the default.

REFERENCES

Akerlof, George A. 1978. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," P. Diamond and M. Rothschild, *Uncertainty in Economics*. Amsterdam, The Netherlands: Elsevier, 235-51.

Behling, Orlando. 1998. "Employee Selection: Will Intelligence and Conscientiousness Do the Job." *The Academy of Management Executive (1993-2005)*, 12, 77-86.

Berry, Christopher M; Melissa L Gruys and Paul R Sackett. 2006. "Educational Attainment as a Proxy for Cognitive Ability in Selection: Effects on Levels of Cognitive Ability and Adverse Impact." *Journal of Applied Psychology*, 91(3), 696-705.

Brogden, Hubert E. 1949. "A New Coefficient: Application to Biserial Correlation and to Estimation of Selective Efficiency." *Psychometrika*, 14(3), 169-82.

____. 1946. "On the Interpretation of the Correlation Coefficient as a Measure of Predictive Efficiency." *Journal of Educational Psychology*, 37(2), 65-76.

Bureau of Labor Statistics. 2016. "Labor Force Statistics from the Current Population Survey – LNS14000000." United States Department of Labor. http://data.bls.gov/timeseries/LNS14000000 (accessed April 15th, 2016).

Chapman, Derek S.; Krista L. Uggerslev; Sarah A. Carroll; Kelly A. Piasentin and David A. Jones. 2005. "Applicant Attraction to Organizations and Job Choice: A Meta-Analytic Review of the Correlates of Recruiting Outcomes." *Journal of Applied Psychology*, 90(5), 928-44.

Converse, Patrick D. and Frederick L. Oswald. 2014. "Thinking Ahead: Assuming Linear Versus Nonlinear Personality-Criterion Relationships in Personnel Selection." *Human*

Performance, 27(1), 61-79.

Dunleavy, Eric M.; Lorin M. Mueller; Ash K. Buonasera; Daniel C. Kuang and Dana Glenn Dunleavy. 2008. "On the Consequences of Frequent Applicants in Adverse Impact Analyses: A Demonstration Study." *International Journal of Selection and Assessment*, 16(4), 333-44.

Fine, Sidney A. and Steven F. Cronshaw. 1999. *Functional Job Analysis: A Foundation for Human Resources Management*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gale, David and Lloyd S. Shapley. 1962. "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*, 69(1), 9.

Gibson, Shanan G.; Robert J. Harvey and Michael L. Harris. 2007. "Holistic Versus Decomposed Ratings of General Dimensions of Work Activity." *Management Research News*, 30(10), 724-34.

Gilster, Paul A.; Barbara Davison and William Dickmeyer. 2001. "Channel the Resume Flood with Applicant Tracking Systems." *The Personnel Journal*, 80(1), 30-31.

Glomb, Theresa M.; John Kammeyer-Mueller and Maria Rotundo. 2004. "Emotional Labor Demands and Compensating Wage Differentials." *Journal of Applied Psychology*, 89(4), 700-14.

Goldberg, Lewis R. 1999. "International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality and Other Individual Differences."

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political economy*, 110(2), 458-74.

Greenwald, Bruce C. 1986. "Adverse Selection in the Labour Market." *The Review of Economic Studies*, 53(3), 325-47.

Grove, William M.; David H. Zald; Boyd S. Lebow; Beth E. Snitz and Chad Nelson. 2000. "Clinical Versus Mechanical Prediction: A Meta-Analysis." *Psychological Assessment*, 12(1), 19-30.

Hambrick, David Z. and Elizabeth J. Meinz. 2011. "Limits on the Predictive Power of Domain-Specific Experience and Knowledge in Skilled Performance." *Current Directions in Psychological Science*, 20(5), 275-79.

Heckman, James J. and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics*, 19(4), 451-64.

Hülsheger, Ute R.; Günter N. W. Maier and Thorsten Stumpp. 2007. "Validity of General

Mental Ability for the Prediction of Job Performance and Training Success in Germany: A Meta-Analysis." *International Journal of Selection and Assessment*, 15(1), 3-18.

Hunter, John E. and Frank L. Schmidt. 1982a. "Fitting People to Jobs: The Impact of Personnel Selection on National Productivity," M. D. Dunnette and E. A. Fleishman, *Human Capability Assessment.* Hillsdale, NJ: Lawrence Erlbaum Associates, 233-92.

____. 1982b. "Fitting People to Jobs: The Impact of Personnel Selection on National Productivity," M. D. Dunnette and E. A. Fleishman, *Human Performance and Productivity.* Hillsdale, NJ: Erlbaum, 233-84.

Hunter, John E.; Frank L. Schmidt and Michael K. Judiesch. 1990. "Individual Differences in Output Variability as a Function of Job Complexity." *Journal of Applied Psychology*, 75(1), 28.

International Monetary Fund. 2016. "World Economic Outlook Database,"

Izyumov, Alexei and John Vahaly. 2015. "Income Shares Revisited." *Review of Income and Wealth*, 61(1), 179-88.

Johnson, Jeff W; Piers Steel; Charles A Scherbaum; Calvin C Hoffman; P Richard Jeanneret and Jeff Foster. 2010. "Validation Is Like Motor Oil: Synthetic Is Better." *Industrial and Organizational Psychology*, 3(3), 305-28.

Judge, Timothy A.; Jessica B. Rodell; Ryan L. Klinger; Lauren S. Simon and Eean R. Crawford. 2013. "Hierarchical Representations of the Five-Factor Model of Personality in Predicting Job Performance: Integrating Three Organizing Frameworks with Two Theoretical Perspectives." *Journal of Applied Psychology*, 98, 875-925.

Leong, Corissa. 2018. "Technology & Recruiting 101: How It Works and Where It’s Going." *Strategic HR Review*, 17(1), 50-52.

Leung, Seung-Ming. 2008. "The Big Five Career Theories," J. A. Athanasou and R. van Esbroeck, *International Handbook of Career Guidance.* New York, NY: Springer, 115-32.

McCloy, Rodney A.; Dan J. Putka and Robert E. Gibby. 2010. "Developing an Online Synthetic Validation Tool." *Industrial and Organizational Psychology*, 3(3), 366-70.

McDaniel, Michael A.; Deborah L. Whetzel; Frank L. Schmidt and Steven D. Maurer. 1994. "The Validity of Employment Interviews: A Comprehensive Review and Meta-Analysis." *Journal of Applied Psychology*, 79(4), 599-616.

Meriac, John P.; Brian J. Hoffman; David J. Woehr and Matthew S. Fleisher. 2008. "Further

Evidence for the Validity of Assessment Center Dimensions: A Meta-Analysis of the Incremental Criterion-Related Validity of Dimension Ratings." *Journal of Applied Psychology*, 93(5), 1042-52.

Meyer, Rustin D.; Reeshad S. Dalal and Silvia Bonaccio. 2009. "A Meta-Analytic Investigation into the Moderating Effects of Situational Strength on the Conscientiousness-Performance Relationship." *Journal of Organizational Behavior*, 30(8), 1077-102.

Meyer, Rustin D.; Reeshad S. Dalal; Irwin J. José; Richard Hermida; Tiffani R. Chen; Ronald P. Vega; Charlie K. Brooks and Vivek P. Khare. 2014. "Measuring Job-Related Situational Strength and Assessing Its Interactive Effects with Personality on Voluntary Work Behavior." *Journal of Management*, 40(4), 1010-41.

Molloy, Janice C and Jay Bryan Barney. 2015. "Who Captures the Value Created with Human Capital? A Market-Based View." *Academy of Management Perspectives*, 29(3), 309-25.

Nye, Christopher D.; Rong Su; James Rounds and Fritz Drasgow. 2017. "Interest Congruence and Performance: Revisiting Recent Meta-Analytic Findings." *Journal of Vocational Behavior*, 98(1), 138-51.

O'Neill, Thomas A. and Piers Steel. 2018. "Weighted Composites of Personality Facets: An Examination of Unit, Rational, and Mechanical Weights." *Journal of Research in Personality*, 73(1), 1-11.

Oh, In-Sue; Seongsu Kim and Chad H. van Iddekinge. 2015. "Taking It to Another Level: Do Personality-Based Human Capital Resources Matter to Firm Performance?" *Journal of Applied Psychology*, 100(3), 935-47.

Ones, Deniz S. and Chockalingam Viswesvaran. 2003. "Job-Specific Applicant Pools and National Norms for Personality Scales: Implications for Range-Restriction Corrections in Validation Research." *Journal of Applied Psychology*, 88(3), 570-77.

Outtz, James L. 2010. *Adverse Impact: Implications for Organizational Staffing and High Stakes Selection*. New York, NY: Taylor & Francis.

Peterson, Norman G.; Michael D. Mumford; Walter C. Borman; P. Richard Jeanneret; Edwin A. Fleishman; Kerry Y. Levin; Michael A. Campion; Melinda S. Mayfield; Frederick P. Morgeson; Kenneth Pearlman, et al. 2001. "Understanding Work Using the Occupational Information Network (O*Net): Implications for Practice and Research." *Personnel Psychology*, 54(2), 451-92.

Phelps, Richard P. 1999. "Education Establishment Bias? A Look at the National Research Council's Critique of Test Utility Studies." *The Industrial-Organizational Psychologist*, 36(4), 37-49.

Ployhart, Robert E. 2012. "Personnel Selection: Ensuring Sustainable Organizational Effectiveness through the Acquisition of Human Capital," S. W. J. Kozlowski, *The Oxford Handbook of Organizational Psychology.* New York, NY: Oxford University Press, 221-46.

Roth, Alvin E. 1984. "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political economy*, 92(6), 991-1016.

Sackett, Paul R. and Daniel J. Ostgaard. 1994. "Job-Specific Applicant Pools and National Norms for Cognitive Ability Tests: Implications for Range Restriction Corrections in Validation Research." *Journal of Applied Psychology*, 79(5), 680-84.

Schmidt, Frank L. and John E. Hunter. 1992. "Development of a Causal Model of Processes Determining Job Performance." *Current Directions in Psychological Science*, 1(3), 89-92.

____. 2004. "General Mental Ability in the World of Work: Occupational Attainment and Job Performance." *Journal of Personality and Social Psychology*, 86(1), 162-73.

____. 1998. "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings." *Psychological bulletin*, 124(2), 262-74.

Stagner, Ross. 1958. "The Gullibility of Personnel Managers." *Personnel Psychology*, 11(3), 347-52.

Steel, Piers; Allen I. Huffcutt and John Kammeyer-Mueller. 2006. "From the Work One Knows the Worker: A Systematic Review of the Challenges, Solutions, and Steps to Creating Synthetic Validity." *International Journal of Selection and Assessment*, 14(1), 16-36.

Steel, Piers; Jeff W. Johnson; P. Richard Jeanneret; Charles A. Scherbaum; Calvin C. Hoffman and Jeff Foster. 2010. "At Sea with Synthetic Validity." *Industrial and Organizational Psychology*, 3(3), 371-83.

Steel, Piers and John Kammeyer-Mueller. 2009. "Using a Meta-Analytic Perspective to Enhance Job Component Validation." *Personnel Psychology*, 62(3), 533-52.

Steiner, Dirk D. 2012. "Personnel Selection across the Globe," N. Schmitt, *The Oxford Handbook of Personnel Assessment and Selection.* New York, NY: Oxford University Press, 740-67.

Stone, Dianna L.; Diana L. Deadrick; Kimberly M. Lukaszewski and Richard Johnson. 2015.

"The Influence of Technology on the Future of Human Resource Management." *Human Resource Management Review*, 25(2), 216-31.

Weiss, Andrew. 1980. "Job Queues and Layoffs in Labor Markets with Flexible Wages." *Journal of Political economy*, 88(3), 526-38.

Wilk, Steffanie L.; Laura Burris Desmarais and Paul R. Sackett. 1995. "Gravitation to Jobs Commensurate with Ability: Longitudinal and Cross-Sectional Tests." *Journal of Applied Psychology*, 80(1), 79-85.

Wonderlic, Eldon F. 2007. "Wonderlic Personnel Test-Revised: Manual," Los Angeles, CA: Western Psychological Services,

Youyou, Wu; Michal Kosinski and David Stillwell. 2015. "Computer-Based Personality Judgments Are More Accurate Than Those Made by Humans." *Proceedings of the National Academy of Sciences*, 112(4), 1036-40.

APPENDIX A: META-ANALYTIC ASSISTED SUBJECTIVE ESTIMATION

To obtain informed conditional estimates of how the validity coefficients of the predictors correlate with the job components, we used the subject matter expert estimation approach of job component correlates, as outlined by Rodney A. McCloy, Dan J. Putka and Robert E. Gibby (2010). We then refined the approach by restricting the experts' estimates to ranges that were viable statistically, considering the known meta-analytic distributions of the validity coefficients from the ρ and sd_ρ values reported by Ute R. Hülsheger, Günter N. W. Maier and Thorsten Stumpp (2007), Timothy A. Judge, Jessica B. Rodell, Ryan L. Klinger, Lauren S. Simon and Eean R. Crawford (2013), and John P. Meriac, Brian J. Hoffman, David J. Woehr and Matthew S. Fleisher (2008). With this we arrived at informed, conditional estimates using *Meta-Analytic Assisted Subjective Estimation*.

For each combination of predictor and job component, we asked the subject matter experts to estimate what the correlation between the predictor variable and job performance would be for a job that is one standard deviation above the mean for the job component. For instance: "For a job that's HIGH (1 SD above the mean) on Things the correlation (r) between performance and GMA will be:". In agreement with Rodney A. McCloy, Dan J. Putka and Robert E. Gibby (2010), we used a variety of stimulus materials. These materials included detailed instructions (involving a video) to train and prepare each expert, the definitions of each job component, and a help menu with answers to frequently asked questions and background material. In addition, for each component, we listed relevant job components (e.g., for Things: "repairing and maintaining equipment", "inspecting equipment, structures, or material", and "controlling machines and processes") and exemplary jobs for which the level of the component is relatively high (e.g., for People: "healthcare practitioners", "management occupations", and "protective service occupations").

To further ascertain the quality of the estimates, we developed a new technique, constraining the estimates to values that could be considered feasible within the ranges found in previous meta-analytic estimates of the mean population correlation (i.e., ρ) and the standard deviations of the correlations around this mean. As depicted in Figure A1, we showed the subject matter experts how their estimated correlation between the job component and predictor ranked in comparison to the meta-analytic distribution of the effect of the predictor on job performance. We then instructed them to adhere to the statistical constraints imposed by this distribution. Specifically, since we asked the experts to provide an estimate for a job that is one standard deviation above the mean for the job component, the estimated correlation between the predictor and job performance was constrained to a range from one standard deviation above to one standard deviation below the mean meta-analytic correlation ρ. If the estimated correlation would be smaller or bigger than $\rho \pm sd_{\rho}$, the correlation corresponding to the moderating effect of the job component on the correlation between the predictor and performance (i.e., the "second-order correlation") would be bigger or smaller than one. A correlation, of course, cannot be bigger or smaller than one. Finally, we further reduced the estimates to 50%, a point at which all models were positive definite, to ascertain the viability of the estimates.

[Insert Figure A1 Here]

Figures

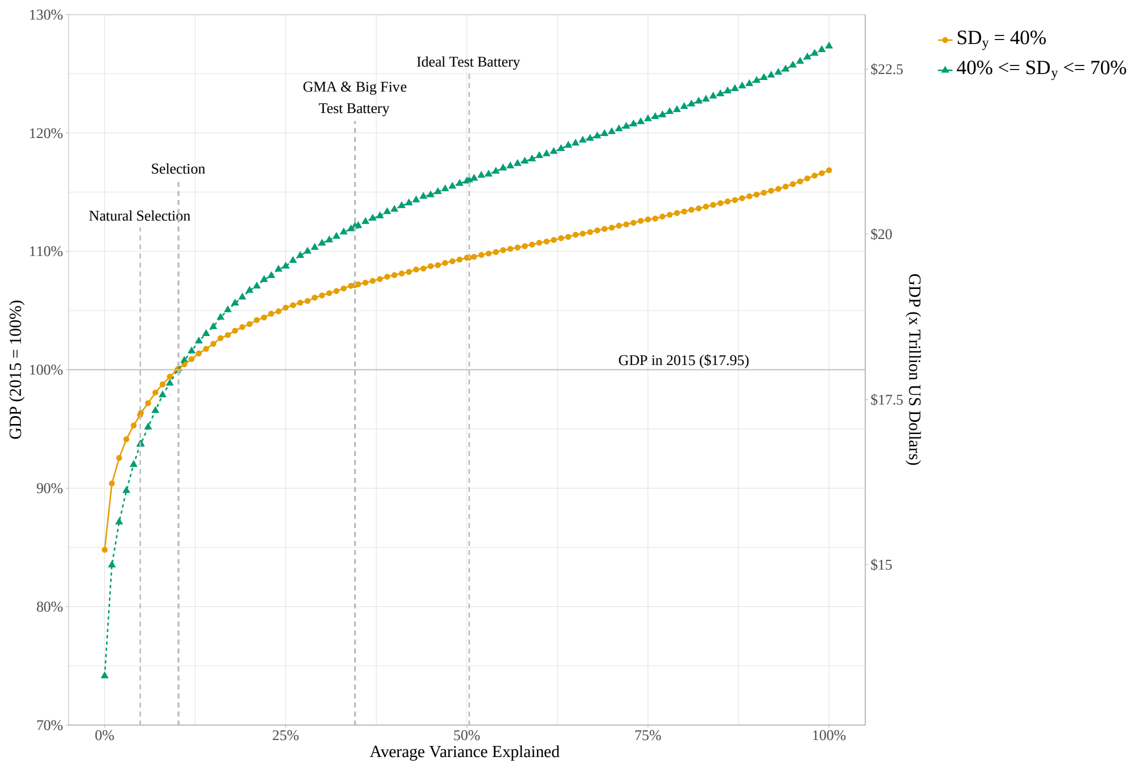

FIGURE 1. EFFECT OF IMPROVED PSYCHOMETRIC ASSESSMENT AND APPLICANT SELECTION ON GDP

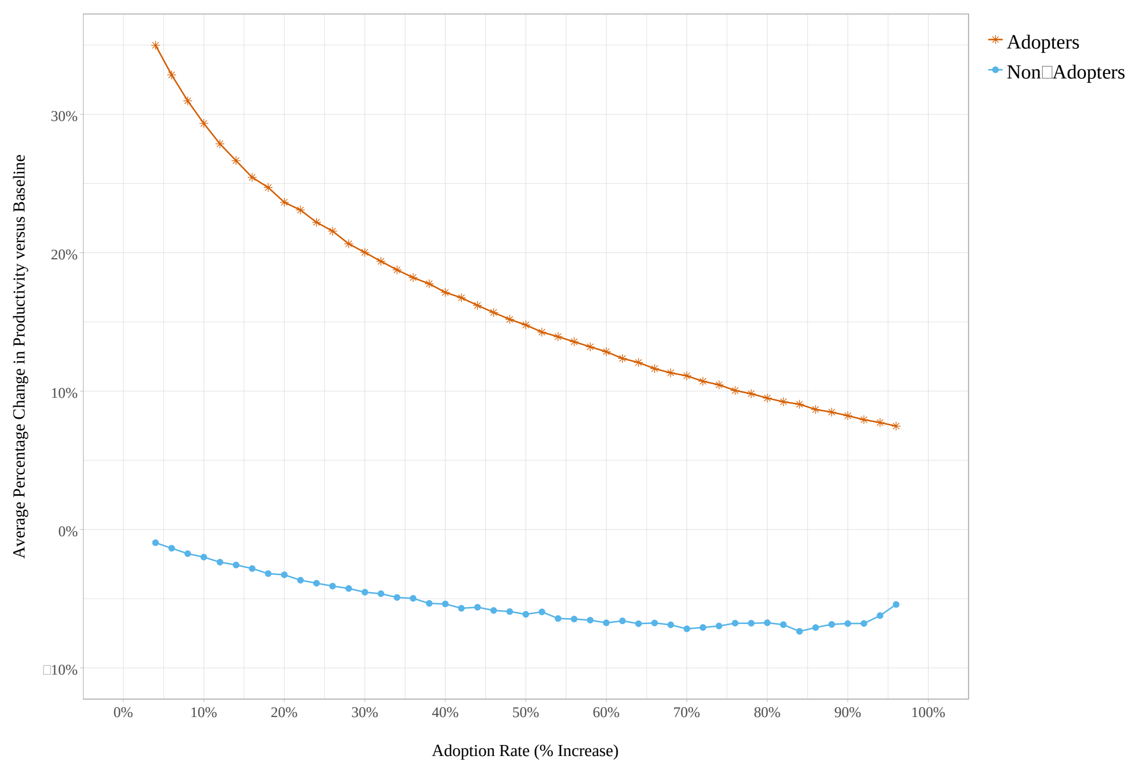

FIGURE 2. MARGINAL RETURNS TO ADOPTION FOR ADOPTERS AND MARGINAL COSTS FOR NON-ADOPTERS

FIGURE A1. OVERVIEW OF KEY INFORMATION AND MAIN STIMULI PROVIDED TO SUBJECT MATTER EXPERTS