Memo

|  |  |
| --- | --- |
| Date: | May 11, 2022 |
| To: | Elizabeth Zachry Rutschow, Ellen Cushing, Lynn Barjorek, Amy Feygin, Christina LiCalsi, Irma Perez-Johnson, and Harry Holzer |
| From: | Christina Curnow, Luke Patterson, Amber Bloomfield, Mason Miller, and Farhan Majid |

# Re: PROMISE Center: The Centers of Excellence at CCC – Labor Market Projections

Goals of PROMISE Center

The PROMISE Center, *Promoting Resilience, Opportunity, and Mobility in Support of Equity*, is committed to uncovering the most effective strategies to build a 21st century workforce development ecosystem that provides equitable access to economic mobility and resilience for millions of American people and the country overall. The PROMISE Center at AIR has established a multi-year partnership with the City Colleges to help them develop, adapt, and strengthen workforce pathways by assessing CCC’s current work and considering whether and how effective sectoral training practices can be adapted and integrated in community college workforce training programs. PROMISE is also working with district and college leaders to advance critical aspects of this pipeline, including how to effectively harness labor market predictions for program planning, build strong employer partnerships, and expand work-based learning opportunities for students. As new strategies and interventions are developed, AIR researchers will seek to document their promise and effectiveness at increasing students’ successful program completion, employment, and earnings, with the goal of replicating the model around the country if it is shown to work

## Overview of the Labor Market Projections Team

Community colleges are one of the pillars of workforce development in the United States. Each year thousands of institutions across the country enroll millions of Americans in academic and workforce training programs intended to help them gain skills to succeed in the workforce. A fundamental challenge facing both job seekers and educators at community colleges is keeping pace with the rapidly fluctuating labor market landscape and the skill requirements of employers. To maximize their students’ chance of success, community colleges must align their curricula to meet these shifting needs. Furthermore, the COVID-19 pandemic has upended labor markets and there is an increasing mismatch between the skills workers offer and those that employers need. For example, in early 2021, 25 percent more workers than previously estimated potentially need to switch occupations due to COVID-19.[[1]](#footnote-2) A 2022 Pulse of the American Worker Survey finds that about a half of US employees are thinking of leaving their job[[2]](#footnote-3). Quit rates had spiked since the pandemic as part of the “Great Resignation.” In Illinois, for instance, the quit rate has increased from 2.0 in Feb 2021 to 2.8 in February 2022.[[3]](#footnote-4)

To meet this challenge, up-to-date and accessible labor market information (LMI) is important to helping job seekers, public policymakers, and educators make informed decisions about a variety of factors affecting their current economic circumstances and futures. For instance, the U.S. Bureau of Labor Statistics (BLS) measures labor market activity and conducts occupational projections in the U.S. economy to support public and private decision-making. Community colleges and educators may use comprehensive LMI data and analysis about local, state, and regional labor markets to determine current job and skill demand, forecast future occupational and skill demand, and inform curriculum design and workforce training programs. Job seekers may use LMI to access jobs in demand, determine the skills required, and find potential employers.

The primary goals for this effort are to (1) help the City Colleges of Chicago (CCC) leadership and the administrators of the CCC Centers of Excellence plan expansions of their current programming, changes to current curricula, and the introduction of new programs; and (2) create a scalable model that can be applied throughout the workforce development ecosystem to enable communities to build economic mobility and resiliency through identifying well paying, in demand jobs and informing programming to help the American workforce gain equitable entry into these jobs. The labor market projections team proposes to acquire and analyze data about future occupation and skill demands in the Chicago area to provide CCC with ideas for potential future programs or changes to current curricula, which we envision being further validated in conjunction with employer engagement conversations.

## Insights from labor market data and the CCC qualitative findings

The findings of AIR’s qualitative team’s interviews with COE leadership and staff indicate that there is strong interest in using labor market data to inform CCC programming. For this purpose, labor market projections will need to be for 2 years or more in the future, as CCC leadership indicated that the COEs will require 2 years’ lead time to effect curricular changes. In addition, the LM projections team understands the need for CCC and its COEs to have relevant, timely information to share with employers in the Chicago area, and that several of the COEs are struggling in building these types of employer relationships. Labor market information could be immediately helpful to informing conversations between COEs and employers. A third application for labor market projections is for marketing purposes: several of the COEs struggle with low enrollment. The qualitative findings indicate that COEs believe that low enrollment may be due to the perceptions potential students have about the kinds of careers associated with their specialty area. For example, the qualitative interview team reported that Daley College cited poor student perceptions of careers in manufacturing as a reason for their low enrollment. Labor market projections could help to provide context around enrollment levels. For example, programs that center around jobs with the highest demand would have the highest potential need for greater enrollment and for occupations where the local demand is lower, the “low” enrollment levels may be aligned with demand. More specifically, labor market analysis that indicates how the demand for skills associated with Daley’s awards might increase in the future could be used in marketing materials to encourage more enrollment. Further, information about the breadth of job opportunities consistent with the skills imparted by Daley’s curricula could also be used to attract more students.

### Labor market data and CCC’s goals

The CCC administrators have EMSI Analyst licenses and use this tool to estimate demand for completers with a given credential. EMSI (which recently merged with BurningGlass, becoming EMSI-BG), is based on job listings data for Cook County, a complete, deduplicated, longitudinal dataset. EMSI Analyst provides users with projections of occupation demand as well as other analyses informed by job listings data and employment and wage indicators from the Bureau of Labor Statistics and Census. CCC leadership and COE administrators have access to the EMSI Analyst tool, but the qualitative research conducted by AIR indicates that this tool is not regularly used by the COE staff. In addition, EMSI Analyst offers projections of occupation demand but does not offer projections of skill demand.

Skill demand is potentially more informative than occupation demand because a specific credential or program may impart skills that map to occupations beyond which the credential or program was originally designed. EMSI-BG has worked with CCC to create a crosswalk of awards to skills, so it will be possible to highlight the relevant awards for any skill-based prediction. Further, skill demand projections can help CCC anticipate how to supplement or shift their current curricula based on how the skills associated with occupations are likely to change. Finally, CCC curricula train students in specific skills. The skill sets required by a whole occupation are often too broad for a single class to cover, so it will be easier for curricula reforms within a COE to target specific skills, rather than an entire occupation. The LMI team proposes to use job listings data from EMSI-BG to project skill demand within the Chicago Metropolitan Statistical Area (MSA). EMSI-BG maps each job listing to the skills relevant for the listing, so it is possible to model the demand for a particular skill over time based on the predicted demand for occupations associated with this skill or the number of job ads mentioning the skill.

To assist CCC and the COE leadership in their employer meetings, the LM projections team can assist CCC administrators in interpreting the occupation projections from EMSI Analyst and the skill demand projections we generate and in translating these predictions into industry-relevant bullet points. The outcome of these analyses and resulting products will be skill-specific projections, mapped to the current CCC programs with an indication of overall projections for each credential (i.e., the percentage of skills associated with that credential that are likely to increase in demand) and an employer information sheet featuring industry-specific information on likely skill demand increases to guide discussion and allow CCC COE administrators to ask employers about their perceived accuracy of the projections.

To address the need of many COEs to increase enrollment, the LM projections team could also provide snapshots of the overlap between the skills associated with an award and the occupations demanding those skills. For example, several programs offered by Daley College may teach the skill Manufacturing Processes. We could provide both the current outlook for the suite of skills taught by each program and the projected outlook for that collection of skills. This information could be used to make programs more attractive to potential students.

Based on our initial conversations with CCC leadership and the qualitative findings from interviews with the COEs, we recommend beginning the LM projection task with analyzing data across the entire Chicago labor market. Then as a second step, we could work more closely with COEs that have an interest in leveraging labor market data in a robust way. This approach will allow us to demonstrate the value of labor market data and identify emerging job growth that may not currently be covered under a COE. This analysis could be presented as part of the COE “reboot” or “reset” discussed by the AIR CCC leadership team during the review of the Qualitative memo. This might have the benefit of allowing some of the COEs to warm up to the idea of receiving assistance related to labor market data (i.e., overcome some of the resistance observed by the qualitative interview team)

## Available data sources

Job postings data will be indispensable in the LM projection team’s ability to forecast labor demand at the occupation and industry level. The team is doing its due diligence in combing though publicly available data that can serve as performance-enhancing covariates to job postings data in forecasting. This section details both public and private data sources that are under consideration for this effort.

The narrative in this memo will focus on our main technical approach, though notes on data series and their potential for additional analyses can be found in the appendices.

### Private Data

Proprietary job postings data provides the unit of analysis for our forecasting model. The best source of these data comes from EMSI-BG. Job *postings* are crucial to our ability to forecast, as they’re the best available proxy for job *openings*, which quantify occupational- and industrial-level labor demand by individual firms. Understanding Chicago’s labor market dynamics at this level will provide CCC a solid empirical basis for their curriculum decisions, which they currently lack. A data set for this purpose is not covered under CCCs existing contract with EMSI.

CCC is planning to purchase an additional data product from EMSI called the Skillabi API. Their use case is to upload syllabi of their numerous courses, at which point EMSI’s proprietary algorithms will map the curriculum to their (also proprietary) skill taxonomy. We plan to work these mappings into our forecasting. This process could prove very valuable if it can identify skills demanded by local employers that aren’t already included in the curriculum.

A potential separate source of useful proprietary data comes from Dun & Bradstreet, which maintains the best-known database of employer and establishment records. These can be of use to CCC in multiple ways: targeting employer partnership outreach, advising potential students with real-world examples of where they could work, and estimating practicality of program-to-job linkages via an understanding of commuting zones. [[4]](#footnote-5)

### Public Data

The data science team is prepared to leverage all publicly available data on both the U.S. population – including details on their economic participation – and the landscape of employers. Broadly, the data series we’re considering come from several sources: DOL Bureau of Labor Statistics (BLS), and a few different Census products – the American Community Survey (ACS), the Current Population Survey (CPS), and the Longitudinal Employer-Household Dynamics project. Together these will help us understand, at a granular level, the population of Chicago and how that population participates in the local economy.

Appendix A lists these public data series, also providing high-level descriptions, notes on their accessibility, and descriptions of how they’re relevant to our task of forecasting labor demand.

## Technical Approach

Forecasting skill demand in a post-COVID world

Forecasting is broadly about using past data to predict the future, and so one of the key assumptions of any forecasting method is that past data is predictive of the future. The efficacy of forecasting is called into question when large shocks disrupt the underlying processes which generate the target outcome and impair the predictive power of data prior to the shock.[[5]](#footnote-6) We are aware of the uncertainty introduced by the COVID pandemic and are using multiple indicators for our forecasting analyses to bolster their efficacy. More information about how forecasting is affected by the pandemic is presented in Appendix C.

Research Questions

We will work with CCC to ensure that their labor market analysis needs are accounted for in our analysis and reporting plan. We plan to investigate the following research questions:

* What is the predicted occupational level demand related to all current COE programs/credentials?
* Is there alignment between occupational demand and enrollment levels?
* What skills map to each occupation/program?
* Are there gaps between the skills taught in the curricula and the predicted future skill demand from employers?
* Is there occupation demand in the Chicago MSA that is not being met through current CCC offerings? If so, for what occupations and what are the credential requirements for those occupations (e.g., Associates degree, certificate, Bachelor’s degree)? What are the salaries for these occupations (i.e., are they good paying jobs?)

After conducting analyses to answer our research questions, we will produce the following for CCC:

Estimated levels of demand for each selected skill 2 years, 4 years, and 6 years into the future. CCC and the individual COEs can use these projections to identify skills that should be included or emphasized in different curricula. The comparison across 2, 4, and 6-year projections can help the CCC prioritize the skills: skills anticipated to grow significantly can be prioritized in the curriculum (provided they are appropriate to the COE’s profession).

Data visualizations, reports, and other technically accessible products for CCC staff to use in conversations with employers, to use in marketing materials, and to inform future curriculum development.

This work will also result in contributions to the field that are discussed next.

## Contributions to the Field

As noted, the job market post-COVID may never completely return to its pre-COVID patterns. Particularly now, there is a strong need for research examining how demand for specific skills has changed and is likely to change in the future. The findings from the LM projections team can contribute to this body of knowledge and can be translated into potential implications for the workforce ecosystem. We anticipate our work advancing the field in several important topic areas:

1. Mapping skills to job opportunities

Skill-based hiring is becoming more common, with fewer job ads listing a required education level[[6]](#footnote-7) The labor market recovery post-COVID has been slower for racial/ethnic minority groups in the US; pre-COVID, members of these groups were also more likely to work at jobs that gave them both an increased risk of contracting COVID and a higher risk of losing their job due to an outbreak.[[7]](#footnote-8) More information about how one’s current skill portfolio compares to market demands will both provide better information for job seeking and for reskilling decisions for members of these groups, particularly given that many racial/ethnic minority groups are more likely than average to have obtained some post-secondary education without completing a degree.[[8]](#footnote-9)

1. Seeking stability in projections

Our approach will allow us to see, among a set of indicators, which are the most reliable predictors of skill demand, and bring to bear methods for overcoming the limitations of online job data using a US dataset (Garasto et al., 2021 looked at job data for the UK only). This will serve to improve the accuracy of labor market projections in future projects, saving time and money for both job seekers and credential providers.

1. Exploring the relationship between skills and awards

It will also be possible for our work to investigate how skills tend to cluster in job listings, providing evidence of both correlated job demands and how employers think about the larger skillset associated with a job. Identifying such skill clusters can also allow us to look at occupational adjacencies to identify reskilling opportunities, especially for occupations that are waning in demand or that do not offer liveable wages, thus contributing to economic resiliency for individuals and communities. Our work could also uncover the skills commonly listed with different types of awards. The frequent co-occurrence of a required certificate or degree and a skill suggests that employers see that skill as highly relevant to work requiring that award, and if the curriculum for that award does not already convey the skill, it would be beneficial if it did.

1. Examining the relative value of skills

Wherever salaries are listed in the job ads, it will be possible for our team to explore the skills, within a particular occupation, that are most strongly associated with salary – a COE completer may be qualified to apply for a job given their skillset, but which skills are likely to make them competitive for the most highly paid positions? This type of information is not only of general interest but could guide COE completers in seeking supplementary training.

Appendixes

Appendix A: Tabulation of Data Sources and Relevance to Forecasting Task

**Private Data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data Source** | **Data Series Name** | **Degree of Need** | **High-Level Description** | **Relevance to Forecasting Labor Demand** |
| EMSI | Job Postings | Must have | Detailed data on job postings, including detailed occupational and industrial codes and specific skills from EMSI’s proprietary taxonomy | Provides the units of analysis for our forecast of labor demand, along with rich detail we can leverage to inform CCC programming and strategy. |
| EMSI | Analyst Tool | Replaced by Job Postings data | Dashboard of real-time job postings data with a UI that allows users to drill down into largely the same set of details available in the raw job postings data. | CCC currently has Analyst seats at each COE, but they are significantly underutilized. After our engagement with CCC ends, administrators can use the Analyst tool to validate our labor demand forecasts with single real-time job posting data points. |
| EMSI | Skillabi API | Helps CCC longer-term and saves work for us | Maps EMSI’s proprietary skill taxonomy to CCC curriculum via syllabi | Allows us to link labor demand estimates and forecasts (via job postings data) to individual CCC programs. |
| EMSI | Career Coach API | Not of current value to EMSI | Aggregates EMSI’s job postings data into estimates and projections at the career level, i.e., groups of related occupations. | Strictly less useful than EMSI’s job posting API. If our forecasting product proves helpful to CCC, they could save money by removing this subscription for the future. |
| Dun & Bradstreet | Employer records | Can help COEs with employer engagement and marketing | Geolocated data on employers and establishments with detailed industrial codes (5-6-digit NAICS), points of contact, and contact information | Identify actual employers and establishments where CCC COE graduates can work, allowing for a more sophisticated analysis of commuting times. Target employer outreach efforts for COEs who would benefit from growing their partnerships. |

**Public Data**

**None of these public data series are strictly necessary for our labor market forecasting analysis. They are listed here because they *may* improve our forecasting model and have potential to help CCC in other ways unique to each data series.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data Source** | **Data Series Name** | **Geographies Available** | **High-Level Description** | **Relevance to PROMISE and Forecasting of Labor Demand** |
| BLS | Occupational Employment Statistics | State, county, MSA | Employment counts and wage statistics by O\*Net occupation | Best source of employment counts and actual wages earned for O\*Net occupations in Cook County |
| BLS | Location Quotient | State, county, MSA | Index of industry concentration relative to U.S. average | Data on which industries are concentrated in Cook County. Starting point for understanding how economic/COVID shocks propagate throughout Cook County, e.g., service sector lockdowns, or how a national housing market bust would affect downstream industries like wood production. |
| BLS | QCEW | County, MSA | “Quarterly Census of Employment and Wages” – counts of employment and wage bins by industry (1-2 digit NAICS) | Creation/destruction of employers/establishments and wage trends by industry. Validate employer claims about stability of labor demand. |
| BLS | Employment Projections | Nationwide | 10-year forecasts of occupational and industrial employment | Can be combined with Location Quotient and QCEW data to project occupational/industrial demand for certain jobs in Chicago – though projections historically haven’t performed well |
| BLS | Industrial-Occupational Matrix | Nationwide | Yearly employment counts for every cell of the 6-digit O\*Net code x 6-digit NAICS code matrix | Can be combined with Location Quotient and QCEW data to estimate Chicago’s employment counts for the most granular combinations of occupations and industries, e.g., number of RNs in hospitals vs. number of RNs in nursing homes. |
| BLS | Input-Output Matrix | Nationwide | Shows the flow of commodities from production through intermediate use by industries (NAICS 6-digit) to purchases by final users | Relates industries to each other in terms of production chains and dependencies, denoted by dollar amounts. Understand how industry shocks propagate through other industries, ultimately affecting employment. |
| BLS | Local Area Unem-ployment  Statistics | State, County, City | Monthly counts of labor force, employment, unemployment, and participation rates | Counts of the number of workers and job seekers in Chicago and how these values have changed at the month level. For example, while the IL unemployment rate is low, labor force participation hasn’t recovered – meaning as many as 200,000 workers may return if good opportunities present themselves. |
| LEHD | Job-to-Job Flows | State | Quarterly statistics on hires and separations with a focus on worker reallocation across employers and whether there was a spell of nonemployment. Sliced by firm characteristics (industry, size, age) and worker demographics (sex by age, sex by education, race/ethnicity)  2-digit NAICS available for IL. | Can be combined with Location Quotient and QCEW data to apportion job flows into Chicago from the state level. Provides granular estimates of labor market churn, which is crucial given the economic turmoil of COVID, global supply chains, inflation, and interest rate hikes. |
| LEHD | Post-Secondary Employment Outcomes | Nationwide | Earnings and employment outcomes of graduates of select post-secondary institutions in the U.S. | Benchmark for earnings and employment outcomes of graduates in similar educational programs to COEs |
| LEHD | Origin-Destination Employment Statistics | Census blocks (maximum granularity) | Detailed data on employment counts and worker commutes, sliced by 2-digit NAICS, 3 earnings bins, age bins, race, gender, educational attainment, firm age, firm size | Understand actual commuting patterns in Chicago to shed light on how practical commuting constraints limit the job prospects of potential students. |
| O\*Net | All dimensions | N/A  (data covered by BLS) | Detailed taxonomy of occupations  Caveat: Generally less accurate and rich than job postings data. | Related Occupations, Career Paths, Tasks, Skills, Knowledge, Abilities, Interests, Detailed Work Activities, Work Context, Experience Requirements, Certifications, Work Styles, Work Values |
| ACS | Estimates | Census blocks (maximum granularity) | 1% sample of U.S. population, chock full of demographic and economic participation information at the smallest available geographies, e.g., Census tracts and blocks. | Counts of potential students and their commuting proximity to COEs at the maximum available geographical granularity. Spatial comparisons of economic/social inequality. |
| ACS | Microdata | PUMAs (unique, with min 100k population) | Microdata (rows of actual people/households) available at PUMA-level | Survey weights allow for production of custom indicators not already made available by the ACS estimates, e.g., number of disconnected non-white Hispanic men aged 16-24 (at the PUMA geography). |
| CPS | Microdata | State, County, MSA | Household-level (short) panel of microdata with demographics and some economic participation data | Provides a representative sense of how demographic and economic participation data changes from month-to-month at the individual and household level. |
| CPS ASEC | Microdata | State | Detailed economic participation microdata as a supplement to the Basic CPS. Panel is even thinner than Basic CPS. | Provides demographic and earnings profiles for certain non-granular industry and occupational codes. Only updated once yearly, in March. |

### Appendix B: Accessibility of Data

EMSI’s suite of data offerings includes a variety of accessibility options which allow us to efficiently integrate them into our analyses while enabling reproducibility. Raw job postings data are available via two sources. Snowflake, a cloud database featuring a UI that allows users to query the data, was recommended by EMSI salespeople. The other option is an API, which is more austere in the window it provides to the data. Either way, the data science team will write Python code to extract and load job postings data into our forecasting model.

All BLS, CPS, ACS, and O\*Net data is available via API, enabling the data science team to pipe data into forecasts that integrate new data as soon as they become available. Notably, LEHD data is only available in direct downloads. These data are published only quarterly, however, meaning that updating our work with the latest LEHD data requires minimal effort.

### Appendix C: Technical Approach Details

#### The Impact of COVID on Labor Market Forecasting

The COVID-19 epidemic had a profound effect to the labor market in a multitude of ways. While the official recession lasted only two months (<https://www.nber.org/news/business-cycle-dating-committee-announcement-july-19-2021>), unemployment has only just returned to pre-COVID levels (<https://www.bls.gov/charts/employment-situation/civilian-unemployment-rate.htm>), and job separations have been steadily growing since the end of the recession (<https://www.bls.gov/charts/job-openings-and-labor-turnover/opening-hire-seps-level.htm>).

This imposes limitations and caveats on our attempts to forecast skill demand that we will keep in mind throughout the project:

* The predictive value of pre-COVID job postings data may be minimal, leaving us with a narrow time frame of post-COVID recession data that may limit the statistical confidence of our predictions.
* The impact of COVID is not over – underlying economic circumstances continued to evolve after the pandemic. This may disrupt even post-COVID recession data’s efficacy at portraying the future.

Several articles cover methods to address these concerns when conducting forecasting in a post-COVID era. We will take these methods into consideration as we develop our forecast models:

* Carriero et al. (2021) propose an alternative approach which combines stochastic time variation in volatility with an outlier correction mechanism.[[9]](#footnote-10)
* Lenza and Primiceri (2020) propose lower weights on COVID observations by allowing for higher volatility of associated residuals.[[10]](#footnote-11)
* Primierci and Tambalotti (2020) discuss an approach to synthetically modeling a COVID shock based on previous historical recessions events, but modifying the propagation of the shock to assume a speedier recovery than historically observed in previous recessions.[[11]](#footnote-12)

#### Modeling Steps

We will apply these methods in following steps to obtain the projections for skill demand:

1. **Data cleaning:**
   1. **Tag skills in job postings.** EMSI uses natural language processing to tag job postings with thousands of skills in their proprietary taxonomy. We will perform quality checks on samples of EMSI data to ensure they are coded properly, and implement corrections based on our findings
   2. **Transform job postings data into time series data.** For each skill, this involves counting the frequency of job postings within each time step most appropriate for forecasting (likely to be monthly) containing that skill.
   3. **Weighting adjustments.** EMSI data only captures online job postings, and not all occupations post the same share of job openings online. Generally, white-collar jobs are more likely to be posted online than blue-collar jobs (<https://www.economicmodeling.com/data/>). Previous literature has investigated ways to re-weight job posting data to account for this disparity[[12]](#footnote-13), and we will investigate weighting our data with such methodologies to account for this.
2. **Train the forecasting model**.
   1. **Identify skills to project.** Not all skills will have sufficient sample size to form projections; by definition these are low-demand skills of less importance to CCC. We will identify and target skills that are within CCC’s capabilities to train for. This will be done by examining the prevalence of skills within CCC’s existing curricula, which have already been coded for skills by CCC’s use of EMSI’s Skillibi tool. We will also use the results produced by the PROMISE center qualitative team to inform the skills to model. This may still leave hundreds of skills of interest for projections. However, there are large economies of scales to model development, and we anticipate the number of skills selected for forecast to have minimal impact on the level of effort required for this task.
   2. **Model tuning.** We will develop the forecasting model based on which calibration performs best on training data. This includes testing a variety of forecasting methodologies, hyperparameters, and feature sets to identify which combination of these best predicts historical job postings demand. This also includes cross-validation of results with different samples of training and test data to minimize the possibility of model overfitting. Model configuration performance may vary for different skills, and we will likely end up with different predictive models for different skills.
3. **Produce forecasted results**. After finalizing model configurations, we will run the forecasting models to produce estimated levels of demand for each selected skill 2 years, 4 years, and 6 years into the future. We will examine results for irregularities/inconsistencies and recalibrate/rerun models as needed until these issues are addressed.
4. **Develop dissemination products**. We will work to develop data visualizations, reports, and other technically accessible products for CCC staff to use in conversations with employers, to use in marketing materials, and to inform future curriculum development.

#### Univariate versus Multivariate Projections

Univariate projections (including methods such as ARIMA) of a given skill’s demand based solely based on past skill demand are not robust to potential shocks (in the case of ARIMA, it violates the assumption of stationary trends[[13]](#footnote-14)). To mitigate this, we intend to utilize multi-variate forecasting methods (such as Vector Auto Regression or ARIMAX) that incorporate not only the past levels of skill demand as predictors, but also other time series of local labor market indicators in Chicago. We will investigate including the following measures to supplement historical skill demand levels from job postings:

* Occupation-level labor demand in occupations related to the skill
  + Occupations related to the skill will be identified using EMSI’s skill taxonomy and empirical observation of the occupation
  + Occupation-level labor demand will be obtained from EMSI projections
* Replacement skill demand relative to skill replacement supply
  + Replacement skill demand is measured by the job postings analysis described above.
  + Replacement skill supply could be inferred from the resume database EMSI has, or from the occupations previously held by the local unemployed population as observed in the CPS.
* Net employment growth
  + Local Chicago employment by industry will be obtained from BLS State and Metro Area Employment, Hours, & Earnings
* Relative real wage growth in a field (relative to similarly educated workers in other occupations or industries)
  + Local Chicago wages by industry will be obtained from BLS State and Metro Area Employment, Hours, & Earnings

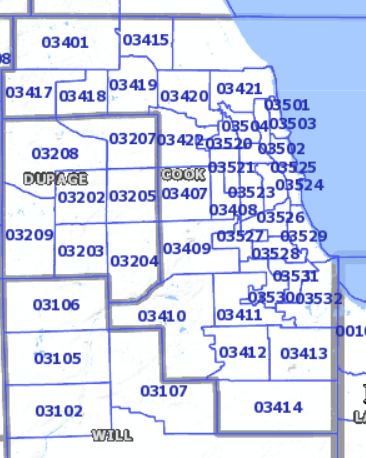
### Appendix D: Notes on Auxiliary Analyses and Utility of Public Data

While many of the public data series described in Appendix A have limited geographical, industrial, or occupational granularity, the data science team can combine data series in certain ways to shed light on labor market dynamics in smaller granularities more relevant to individual COEs.

For example, LEHD ODES data provides details on commuting patterns at the Census block level – the smallest geography available in U.S. government data. Additionally, it tells us which industries the workers in any Census block commute to and from, though only at a high level (2-digit NAICS codes). The data science team can provide a sense of which occupations these workers hold by leveraging separate data series from BLS, namely the location quotient and occupational-industrial employment matrix.

As a first step, the location quotient allows us to make data-driven estimates of additional granularity (extra digits) to the NAICS codes already in the LEHD data. From there, the occupational-industrial employment matrix provides a rough crosswalk to the occupations the workers in the LEHD data may hold. If the data provide sufficient signal – which is liable to vary by industry and occupation – we can recover the real-life commuting patterns for any occupation or career path of concern to CCC.

Administrators and advisors can use these data in a number of ways: to assess commuting constraints of the career paths they train, to provide this information as guidance to students, to use it to inform marketing or employer outreach efforts, and many more uses that will become clear as we engage further with CCC.

Exhibit 1: Overlay of PUMAs (blue regions) with Cook County, IL (grey outline)

1. <https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-after-covid-19> [↑](#footnote-ref-2)
2. https://www.weforum.org/agenda/2022/03/workers-jobs-employee-business-reshuffle/ [↑](#footnote-ref-3)
3. https://www.bls.gov/news.release/jltst.t04.htm [↑](#footnote-ref-4)
4. The price of these data was not included in earlier estimates. Pricing of these data takes a similar structure to job postings data. A “seat” option similar to the EMSI Analyst tool allows for individual querying of employer records with a user interface, and credits can be purchased for exports of these records at a few cents per record. A year’s subscription runs in the low 5-figures, with different quality options available. Another alternative is a one-time data pull with highly variable cost, but this option is not likely to suit CCC’s needs. [↑](#footnote-ref-5)
5. Bobeica, Elena, et al. (2021). "After floods and pandemics: How to obtain a meaningful forecast." [↑](#footnote-ref-6)
6. Fuller, Langer, and Sigelman. *Skill-based hiring is on the rise*. Harvard Business Review. Retrieved from: <https://hbr.org/2022/02/skills-based-hiring-is-on-the-rise>, 2022. [↑](#footnote-ref-7)
7. Organisation for Economic Co-operation and Development.(2022). *The unequal impact of COVID-19: A spotlight on frontline workers, migrants and racial/ethnic minorities*. Paris: Secretary-General of the OECD. Retrieved from <https://www.oecd.org/coronavirus/policy-responses/the-unequal-impact-of-covid-19-a-spotlight-on-frontline-workers-migrants-and-racial-ethnic-minorities-f36e931e/> [↑](#footnote-ref-8)
8. National Center for Education Statistics. *Status and Trends in the Education of Racial and Ethnic Groups*. (2019). Retrieved from <https://nces.ed.gov/programs/raceindicators/indicator_rfa.asp>. [↑](#footnote-ref-9)
9. Carriero, A, T E Clark, M Marcellino and E Mertens (2021), “Addressing COVID-19 outliers in BVARs with stochastic volatility”, Working Paper 202102R, Federal Reserve Bank of Cleveland, revised 09 Aug 2021. [↑](#footnote-ref-10)
10. Lenza, M and G E Primiceri (2020), “How to estimate a VAR after March 2020”, NBER Working Paper 27771. [↑](#footnote-ref-11)
11. Primiceri, Giorgio E., and Andrea Tambalotti. "Macroeconomic Forecasting in the Time of COVID-19." Manuscript, Northwestern University (2020): 1-23. [↑](#footnote-ref-12)
12. Garasto, Stef, et al. Developing experimental estimates of regional skill demand. ESCoE Discussion Paper 2021-02). London: Economic Statistics Centre of Excellence and the Office for National Statistics. Retrieved from: https://www. escoe. ac. uk/boosting-labour-market-intelligence-for-local-decision-makers, 2021. [↑](#footnote-ref-13)
13. Zhang, Mingda. "Time Series: Autoregressive models AR, MA, ARMA, ARIMA." University of Pittsburgh (2018). [↑](#footnote-ref-14)