

**October 23, 2019  
1:30 pm ET**

Coordinator: Welcome and thank you for standing by. At this time all participants are in a listen-only mode until the question-and-answer session of today's conference. At that time you may press star then the number 1 on your phone to ask a question. I would like to inform all parties that today's conference is being recorded. If you have any objections, you may disconnect at this time. I would now like to turn the conference over to Earlene Dowell. Thank you, you may begin.

Earlene Dowell: Thank you (Jennifer) and thank you Greg Pewett from the U.S. Census Bureau for hosting our Webinar today. On behalf of the U.S. Census Bureau and a partnership with the Council for Community and Economic Research and the Labor Market Information Institute, welcome to our October LED Webinar.

It is with great pleasure that I introduce my friend and colleague, Matthew Graham as he presents recent updates to LEHD origin destinations employment statistics for OnTheMap. In August the LEHD released two years -- 2016 and 2017 -- of data for the LODES datasets.

These data were also added to the OnTheMap Web application. This Webinar will discuss the recent data update, associated application changes, and work through a series of demonstrations using OnTheMap and the LODES to answer specific questions about where workers are employed and where they live, and the relationship between the two.

Matthew Graham is the Lead of the development and application innovation group within the LEHD program in the Center for Economic Studies at the U.S. Census Bureau.

Over the last decade, he has led teams to develop and implement new confidentiality protection systems, new public use datasets such as LODES that use those privacy mechanisms and Web-based data dissemination exploration tools such as OnTheMap which was a winner in 2010 of the U.S. Department of Commerce's gold medal for scientific engineering achievement.

Matt has an MA in Urban Planning from UCLA as well as an MS in Mechanical Engineering and a BS in Physics from MIT. With that, I welcome Matt.

Matthew Graham: Thanks, Earlene. Again my name is Matthew Graham. I'm here to talk about the recent updates to LODES, that is the LEHD Origin Destination Employment Statistics, and our Web application that distributes LODES which is called OnTheMap.

I do just want to say quickly that all of the data I will discuss today is public use data. I will not be discussing any confidential data and I will be speaking really only about LODES, the LEHD program and the products and tools that we produce.

This is a rough agenda for what we will be doing today. The times that are listed are estimates. I will be taking questions. I encourage you during this Webinar to chat your questions in using the chat feature in the Webinar.

I will pause at several points to read and answer those questions as best as possible. At the end of the session, if there is time we will open-up the phone line and take questions via phone but again please do feel free to chat your questions in and likely that I'll get to them more quickly that way.

So I'll do a quick overview of the LEHD program today, then I'll spend most of the time talking about LODES and OnTheMap. Okay, the goal with this section is to give us all a baseline. I see that there are some long-time friends of the LEHD program and people who are relatively knowledgeable about our data and tools on the Webinar today.

But I'm also guessing there are folks out there who are not as familiar with the program and our data and tools and so I'll start by addressing some of the basic topics. So, what are LEHD and LED? So, LEHD, and there will not be a test on this particular point, LEHD stands for Longitudinal Employer-Household Dynamics.

It is the program within the Census Bureau and it connects administrative records with Census and survey data to produce new data products without significant additional respondent burden. LED is the Local Employment Dynamics Partnership. It is through that partnership that the LEHD program gets access to significant sets of administrative records on workers and firms to produce the data that we then release to the public.

So in a pictorial sense, I wanted to give people an idea of how we do this so LEHD works mainly in the areas of workers, jobs, and businesses and so we collect quite a few different administrative Census and survey data on these different things, jobs, businesses and workers.

And this is a visual representation of that so firm data we get from a number of sources. One of those is the Quarterly Census of Employment and Wages, QCEW that comes to us through our state partnership. We also get information from economic survey data as well as the Business Register which is a Census Bureau internal product that is sourced out of tax data and other sources as well.

Jobs data come to us from two main sources. The first is unemployment insurance wage records that come to us again from our state partners. These cover private employment and local and state government employment and the coverage includes all those workers who are covered by unemployment insurance within the state.

That specific coverage can vary from state to state but in general what we're talking about is private employment as well as state and local government. We also get federal employment data from the Office of Personnel Management OPM and I will discuss a little bit more about that specific source later in this talk.

Finally we get person data from a couple of different sources. These include federal records so from different programs, different agencies we get access to administrative data on people and we also have access to the demographic Census and survey such as from the Decennial Census or the American Community Survey.

We bring all of that together into a linked national jobs dataset so what we're really focused on here is jobs. We're looking to understand employment and jobs in the United States and produce data products out of that.

In order to produce those data products, we employ cutting-edge confidentiality protection mechanisms. Once we have the Linked National Jobs Data, we can't release it as is, it is confidential and so we protect that confidentiality and through that produce public use data products.

In general the coverage for this infrastructure is about 97% of private employment and most state, local, and federal jobs. Right now the data availability is 1990 through 2018 although that may now be 2019 as we're starting to roll out 2019 data for some of our data products.

Okay, the data products that we produce here in LEHD there are four main ones and I'll just highlight a few features of those so the first is Quarterly Workforce Indicators. This is a data product that includes 32 indicators on the workforce.

It includes things like employment, hirings, separations, turnovers, that kind of thing. It covers about 150 million job records or more that are processed each quarter and for some states the data series starts in 1990 although for other states, the data series is much shorter because that was the available history that we were able to access.

The next data product is LODES, the LEHD origin destination employment statistics. That will be the main data product we'll be talking about today, but in a nutshell it connects a job or a worker's employment location to their residential location. It's at the census block detail. It has less characteristic detail than the Quarterly Workforce Indicators but much, much, more spatial detail.

The third data product that we produce is called Job-to-Job Flows. This is also an origin destination data product, but it's one that connects jobs. When

you leave a job, where do you go? Do you go to another job or do you go to a spell of non-employment? And when you change jobs, do you also change industries? Do you change geography?

Those are the kinds of questions that Job-to-Job Flows can answer. And, finally our most recent data product is Post-Secondary Employment Outcomes, PSEO. This is a pilot project right now. We have these data for a small handful of university systems but it looks at the question of if you get a degree from this university at 1, 5, and 10 years, what are your employment outcomes? What do your earnings look like and where are you working, what industry are you working in?

This is another way to think about the data products and I will say at this point that these slides along with the recording of this Webinar and a transcript will be made available I believe about a week after this Webinar is finished so you will have access to these slides for reference purposes.

But this slide here shows just a different way of saying what I just described about each of the data products and it's really a question of what do you want and this lets you know which data product might be able to help you.

LEHD also produces a series of dissemination tools and applications on the Internet so these are all Web-based tools that anyone can access. They're free and available to all and they make the data products that we produce available for exploration as well as download.

So the tools I'll mention just quickly, J2J Explorer, OnTheMap, we'll be talking about OnTheMap much more today. OnTheMap for Emergency Management, QWI Explorer, the LED Extraction Tool, and the PSEO Data

Viz which again is a pilot program, and each of the Web sites is listed in this slide as well.

Here's another way of understanding the different access points for the LEHD datasets. So, if you want for example QWI data, you have a number of places to go to get access to that data and I will note that all of the text here that is in blue and underlined are links to the places you can go so again when you get access to this slide deck, you can click on these things and it will take you, your browser will take you there.

So QWI can be accessed via QWI Explorer. If you want to explore the data, answer some basic questions or get some basic visualizations of the QWI data. If instead you prefer to download the QWI data in bulk so that you can put it through your own software or your own analytical process, you can get access to that data in two forms.

One is through our raw data download where we just give you the so-called flat files. They're just CSVs comma-separated variable files and you can download them in whole or in part and then you can also access the raw data through the LED Extraction Tool.

Finally QWI is also available through the Census Bureau's API. I will note for those of you familiar with the API and who have already started using QWI through the API, we have been having a glitch with the API and QWI the last week or two.

My understanding and hope is that that glitch will be resolved today if it's not already resolved. It's my understanding that the issue is now understood by the API staff and that will be fixed shortly. But this is, the API is a way to get

access to the data if you're building your own Web applications and need to make live queries to the data.

The other datasets, LODES, J2J, and PSEO all have exploration applications so those are OnTheMap, OnTheMap for Emergency Management, Job to Job Explorer, J2J Explorer, the PSEO Visualization Tool and you can also get all of those datasets through Raw data download.

We are working to develop API access to those other datasets for the future. Okay, I'm going to pause and see if there are any questions so far about what I've spoken on. So there is a question on LODES and it's how is home-based employment treated in LODES? I'm going to hold that question for just a minute because we will get there.

That's a great question and I'm just going to hold-off and see if there are any other general questions. Oh, this is great. Are employees at religious institutions included in the LODES dataset? I'm actually going to expand that question a bit and say are employees at religious institutions included in any of the LEHD datasets? The answer there is maybe.

It depends if that religious institution is covered under a state's unemployment insurance laws. If for example the religious institution every quarter pays into the state unemployment insurance system for that employee, then the answer is yes. We will have information on them there. I'll talk a bit more as we get into the definition of what the LODES dataset is.

About exactly what kind of information we need to see to make it into LODES but so that's really the test is are you covered by unemployment insurance law? If you for example were fired or let go, or laid off, would you

be able to go to your state to claim unemployment insurance? That's the question.

Let's see, if a person works two part-time jobs, are they counted twice in the LEHD data? I'll just expand from LODES and the answer is potentially yes. We count jobs and I'll discuss what qualifies a job but if you are in the data twice, then we see you twice.

In LODES we have some specific definitions for job that help you as a user determine whether you want to see those additional jobs and I'll talk about those in a minute. So there's a question here, what tool would be good when trying to determine the age demographics in specific counties?

Okay, so when if the geographic entity is a county and you're looking for age distribution and you're focused on workers, then you should be looking at the Quarterly Workforce Indicators, QWI. Those have data at the county level. They have relatively good age distributions, they're not as good as for example what you'll see in the decennial data, or ACS, but again this is focused on workers.

If you're interested in the demographics of the population as a whole, you'll want to look at datasets that the Census Bureau produces about the whole population so again for example decennial or the American Community Survey.

Is, let's see here, is Census Business Builder also being integrated with OnTheMap? Census Business Builder is another tool that is produced by the U.S. Census Bureau. We currently do not have plans to integrate those two applications. That said, there is a long-term strategy within the Census

Bureau to bring together lots of the functionality of the tools we already have and to integrate that functionality widely.

So maybe in the long run, functionality that you see in Census Business Builder and the LEHD applications will be put together. All right, I'm going to pause there on answering questions for now and I'm going to move-on to LODES because I know there are questions specifically about LODES that I've already deferred a bit and I'd like to answer them now.

Okay, the first thing I want to talk about is the LODES update. This will be very brief but for some of you it may be the most important part of this presentation so the big news and Earlene mentioned this at the beginning of the session is that LODES for 2016 and 2017 were released at the end of August this year.

There had been a delay on 2016 and 2017 data. We were able to put them out at the same time. Both years are now available in OnTheMap and in the raw data files. Additionally we backfilled Wyoming data for 2014 and 2015. Occasionally because of delays in signing or extending or renewing agreements with states, we do see lapses in state data.

When we see that, we eventually, well, we are hopeful that we will eventually be able to fill-in that data and that was the case with Wyoming. There was not data for Wyoming in 2014 and 2015, the last few years but with this release in August, we were able to go back and fill that in. And I'll come back to that point in one second. The other thing we did with this data release is update to TIGER 2018.

For those of you who are not familiar with TIGER, it is the base geography set that the Census Bureau uses and the Census Bureau keeps that up-to-date to

keep in line with current political and legal boundaries. And so that's what we have done with this latest release is update to TIGER 2018 to get the most recent political and legal boundaries as the Census Bureau knows them.

If the Census Bureau doesn't know about them, for example, there's a recent change in the boundary of a city, then that is something that will come later when one of the geography division's processes picks that up and puts it into TIGER.

A few other key points, data for Alaska and South Dakota are not currently available in 2017. Hopefully in the future we will be able to backfill those state year combinations. We also do not have right now federal workers in 2016 and 2017. I will talk about that in just another second and because we do not have the federal workers in for 2016 and 2017.

We made a choice to change the default job type of OnTheMap to private primary jobs. Previously and this will make more sense for those of you not super familiar with OnTheMap when I do some demos. Previously the default had been all primary jobs which included federal employment.

However, we did not want users to come to the application and do the analyses they were used to doing but not know that federal jobs were not available and see dramatic changes in their numbers and so we decided to move to the default of private primary jobs so this includes only jobs in the private sector. And I'll point this out when we do the demos in OnTheMap.

So federal jobs are not available again because we are working to renew our agreement with the Office of Personnel Management. However, we didn't have the data in hand and we knew that we were delayed in releasing it and

we finally decided to go ahead and release LODES without federal employment.

As soon as we renew the agreement with OPM, get the data and process it, we will backfill that data into LODES and OnTheMap so that will appear there as soon as we are able to access the raw data and process it and get it out the door but right now federal data is not available for 2016 and 2017.

Okay, the next couple of slides hopefully will answer some more questions people have about LODES. It was developed in the mid-2000s to get at spatially-detailed commuting-like information based out of the LEHD infrastructure, again the dataset based on administrative records largely.

It provides a number of geographic patterns of jobs so one by the employment location so we can see the distribution of employment down to the block level. Two, by residential location we can see the distribution of workers' residences at the block level and then most importantly here is we can see the connections between the two.

So we can ask questions like for workers who are employed in this location, what is their residential pattern? Where do they come from when they get up in the morning and go to work we suspect where do they leave from? Where do they arrive at to do their jobs. And I'll discuss more of that as we work through the application.

These data are tabulated by several categorical variables, age, earnings and industry, sex, race, ethnicity and education, firm age and firm size, ownership of the business, job dominance and job type. So, ownership is for us is private sector industry, a local or state government entity or a federal entity.

Those are three different ownership classes we express in LODES. The job dominance gets back to this question about whether or not more than one job for a person appears in the data. As I said the answer is yes, we do see more than one job. If a person works two covered jobs, we will see them.

If a person works three UI-covered jobs, we will see them. To help folks who are interested more in people and workers than a count of jobs, we have created this sense of job dominance or job primacy and that is simply defined as the job that earned the worker the most income during the reference period.

And at this point it's good to mention the reference period for LODES is Quarter 2 of each year so that's April through June. And then job type is a combination that we use in LODES and OnTheMap of ownership and job dominance so it's just a simple crossing of ownership and job dominance so you're not selecting them independent from each other.

They're mixed together in what we call our job type so job dominance is either all jobs or primary/dominant job. We use primary and dominant interchangeably. All of these data, all of the raw data files are available at this location and that link is embedded elsewhere in this presentation. So a few more details about LODES before we start exploring the data.

Employment for LODES purposes is through our beginning of quarter job definition. When we get data from businesses about their employees, we get information such as who the worker is and how much earnings did they receive during the quarter?

If we see a person has positive earnings that is greater than zero earnings from a business in a reference quarter which is Quarter 2 and we see they also have

earnings from the same business in Quarter 1, then we make the inference that they were employed at that business on the first day of the 2nd Quarter.

So at the beginning of Quarter 2 because we see them having earnings last quarter and this quarter. We say it's reasonable to think that they were employed at the beginning of this quarter and that's what we call a beginning-of-quarter job.

And that's all that LODES covers. The quarterly workforce indicators covers a wide variety of different job definitions but LODES only covers a single job definition, those beginning-of-quarter jobs. I already mentioned the job type, a cross of ownership and job dominance. We have the sense of labor market segments. These are 10 categories.

There are three earnings categories, three industry categories, three worker age, and a total. The origin-destination and the residence and workplace margins can be tabulated by these 10 different categories. We also have some additional characteristics: three earnings, 20 industry, the NAICS sectors, three worker age, two worker sex, six worker race, two worker ethnicity, five firm age, five firm size, and one total.

These characteristics are only available on the residence or workplace margins. They're not available on the origin-destination data. All right, this slide discusses how the raw data files are broken-up. I'm not going to spend really too much time on this at all except to say that the files are broken-up by residence margin, workplace margin and origin-destination.

We also provide some geography files as well that are not mentioned here. Okay, last couple bits on LODES and then we'll get into OnTheMap. If

you're looking to make sense of all these different coverage questions and when if federal jobs available, what variables are available, when, what states?

This slide is a quick reference guide for that so you can see on the left the years of the data so from 2002 when we first started producing LODES through the 2017 data, the number of states that are available in those years, which states are not available during that time, whether or not federal jobs are available, whether or not the race, ethnicity, education and sex variables are available, and whether or not firm age and firm size variables are available.

Right now the heart of the dataset where everything is available is 2011 through 2015. All states are in and all the variables are available for these five years. Okay, and then this slide gets at how to understand the different crossings, when you have OD data, that means you get a residence and a workplace location that you get labor market segments but you don't get any characteristics.

Similar arrangement for the residence margins and the workplace margins. Okay, and then what is OnTheMap, quickly OnTheMap is a Web application that lets you play around with LODES and ask questions and get answers. I can talk about OnTheMap but it's better to show you so that's what we'll do next.

So here is the OnTheMap application. You can get to it in a number of ways. One of them is by doing a Google search for OnTheMap or if depending on, you know, which search engine you may want to type in OnTheMap Census. Either way, you should end-up at [onthemap.ces.census.gov](http://onthemap.ces.census.gov) and your browser should look roughly like this.

I will note that if you are using an older browser specifically, IE, an older version of Internet Explorer, you may have some challenges with OnTheMap. We suggest that you switch to a newer browser, if possible, either Microsoft Edge, Chrome or Firefox all seem to work pretty well.

So I'm going to start-out just by doing a few analyses. I'm going to show you around the application a bit but I will point-out that there is significant help and documentation about using OnTheMap. You can always access that by clicking on the Help and Documentation link at the top right.

And I have assume additional links embedded in the slides that I'll show you at the end of this presentation. So I'm going to search for an area that I want to do an analysis about. I'm going to do an analysis on a place named Sparks. If you're, there are several Sparks. I'm going to choose the one in Nevada.

If you're not familiar with this part of the country, this is the city right next door to Reno, Nevada, in Northern Nevada. I'll move my mouse my map around a bit. Here is Reno, Sparks is over here and I'll back-out just to give you all a sense of roughly where it is.

There is Lake Tahoe, Carson City is down here and then the Bay Area is to the west so I'll zoom back in and then we'll start working our analyses so the first thing I did was I searched for the name of something I was interested in.

I found it in my list of geography here and there's lots of different geographies that are supported and then I clicked on Sparks City, Nevada. That zoomed the map to where I am now and to move forward I just click on this link that says perform analysis on selection area.

That brings me to my analysis settings window. This is where you get to choose everything that will determine your analysis. So the first question I ask is do I want to consider the jobs where the work area is in Sparks or the home location is in Sparks because remember a job has two parts to it.

We've got a worker who has a residence and we have a business that has an employment location and the job connects the two and so we can from a geography perspective we can start by considering either side of this. For now I'm going to leave it on the default work. We'll focus-on the jobs where the employment location is in Sparks.

Then I get this panel analysis type. I'm not going to talk too much about this right now except to say I'm going to leave it on default, area profile. We'll come back and I'll explain a bit more about this. Essentially though it tells us what information we're going to see at the end of this path of the data.

Then I get to choose what year I'm looking at. I'm going to look at the most recent year 2017 and I also get to choose which job type. As I mentioned before, private primary jobs is the default job type. We used to have default job type set to primary jobs which it does include federal employment.

For now we're going to leave it at private primary jobs and only look at the private sector. I'm going to go ahead and click go. After a few seconds, the window will fill-in and we will see employment patterns within Sparks, Nevada.

So I'm going to turn, there are two different visual representations of this. I going to zoom-in just a bit here and you can see a little better so we have what is called a thermal overlay. This is like a density so job density in effect and

we also have a point overlay. I turned-off the thermal overlay. I use it in different situations but for right now I'm turning it off.

And what we see are a series of blue dots, all within the boundaries of Sparks. These dots each represent one census block and the number of jobs where the employment location is in that census block is given by the size and coloring and is shown over here in the legend on the left-hand side.

So we can see there's quite a lot of employment at the southern end of Sparks along Interstate 80. There is other employment distributed throughout Sparks but it's not this dominant job center that is here along Route 80. On the right-hand side we see some notional charts by age, earnings, industry sector, and race.

And we also see some tables and we can scroll through these tables. First we see that the total employment in private primary jobs is about 37,700 in 2017. We can see an age breakdown of that employment. We can see an earnings breakdown of employment and then we can see an industry breakdown.

These are by NAICS industry sectors, and I'll just point-out some of the dominant sectors in Sparks: construction, manufacturing, wholesale trade, resale trade, transportation and warehousing, and also accommodation and food services.

So one thing that's quite neat about OnTheMap is we can dig-in a little deeper here. If I say well there's almost 12% of employment in Sparks is accommodation and food services, where is that located? I can click on accommodation and food services. The map will update and we can see that there are some dominant locations here along Route 80 and then there are other locations distributed throughout.

These are particularly interesting to me and I do know a little bit about Sparks. That's where my wife grew up, because there is a big, there's a cluster of casinos here. There's one really big casino and then there's some other casinos and casinos tend to fall into accommodation and food services. That's not always the case but that's in terms of an industry classification, that tends to be where they fall.

And I do know there are other food and restaurant options and hotel options around Sparks. So, that's some of it but I seem to be missing a lot of the employment that was going-on at the very southern end, so I can go looking for that and there's an interesting cluster here, transportation, retail trade, wholesale trade, and manufacturing, construction.

Let's take a look at that so there we see a bunch, transportation and warehousing and I know just from reading the papers that there's a lot of transportation and warehousing employment in the Reno area in general which is useful for serving the Bay Area in California. It's straight down Route 80 a few hours and so this is actually not too surprising to me.

This is a relatively light industrial area here. It's near a railyard, the railway comes through here and Interstate 80 comes through here so it's not surprising to see transportation and warehousing jobs as well as wholesale trade also clustered in this area so none of this is terribly surprising for me.

So now what I'm going to do is go back and look at a different aspect of employment in and around Sparks. So the first thing I'll do is go down to change settings. That brings me back to the analysis settings panel and what I'm going to do is I'm going to keep the same set of workers, people who are

employed in Sparks in 2017 private primary job but now I'm going to do a different analysis.

I'm going to do the distance direction analysis. This will show me the other side of the origin-destination relationship. So we had been looking at the employment distribution. Now for this same set of workers we're going to be looking at the residence distribution. So, I hit go.

That will spin for just a minute and then it will bring the map back and show me the residence distribution for workers employed in Sparks and there we are. I'm going to turn-off the points this time. I prefer to look at this in the thermal view and I'm going to back-out just a bit.

So I can see the greater Reno area and so now we can see that people who work in Sparks also do seem to live in Sparks but they live throughout the Reno area as well as far down as Carson out to Fernley, other parts of the Washoe Valley, right, and we have over here a distribution of those workers by the distance between their home block and their work block, right?

So we see that over 70% live within 10 miles of their employment location. Another 12% 10 to 24 miles and then there's about 10% here who are listed as greater than 50 miles and I know we've already received a question about this and I'll use this opportunity to talk about it and the question is do people really drive more than 50 miles to their work.

The answer might be yes but the answer also might be no. We have employment location from the businesses which is not always the same thing as a work site. If you telework your employment location is where your employer is but your work site is probably your home.

We do not have information on work sites. We have to make an inference based on the information we've been given from the businesses. This can also be an issue for people who work in jobs that are transitory in their location.

So construction workers, if they're working at a different job site every day, we are not able to adequately capture their work location but we do probably have the main office of their employment or drivers who are on the road all day, clearly their work site is wherever their vehicle is and just in general, long-distance relationships that people may have with their employers.

I had a friend who lived on the East Coast and was technically employed on the West Coast. She did not make that commute every day. That was just a relationship with her employer and in fact her work site was she was a sales person and so her work site was even variable during the day.

So this may capture that sort of relationship. It may also capture cases in which businesses are not reporting all of their work sites. That also happens although it's very hard for us to detect those things a priori.

The last thing I want to point-out is this diagram up here. This is what we call our radar chart. This is a distribution of distance and direction and then we wrap it around a compass rose so the darkest wedges are the closest distances and the size of the wedge represents how many people work for example live to the north less than 10 miles from their work location.

So this gives you a spatial sense of where these workers are coming from so they tend to live to the north, to the west, somewhat to the south and southwest but not really to the east very much. Mostly once you head out of Sparks, you are in the desert and there are a few towns along Route 80 but not

a lot so it's not surprising to see that there's very little to the east and southeast.

We can go back and for comparison we can flip this around and ask about a different set of workers. By changing from work to home, now I'm going to look at workers who live in Sparks but work somewhere else so these are the residents of Sparks and we're going to look at the distribution of their work locations.

And now I'll direct you back to the radar chart because we get a very different distribution. People who live in Sparks dominantly work to the southwest and south of their residence locations, okay? Dominantly, I mean, there's almost nothing to the north and very little to the west or east or the southeast, right?

And that's because most of the employment in Sparks was at the very southern edge and employment outside of Sparks tends to be in Reno or south in the Washoe Valley all the way down to Carson and so folks living in Sparks tend to travel to the southwest or the south when they go to work.

We can see this small pie wedge here to the southeast. If I zoom the map out a bit more and moved it a bit, you could see that to the southeast of Sparks is Las Vegas so I suspect that there are some long-distance relationships between people living in Sparks and having some kind of employment representation in Las Vegas.

The last analysis I'd like to do quickly before I go to questions is the inflow-outflow analysis and this will pull together both of these perspective in one view so it will ask the question of who's working and who's living in Sparks and how are those things related? And so what we see is that about 27,000

workers are employed in Sparks but live outside the selection area, outside of Sparks.

Thirty thousand workers live in Sparks but are employed somewhere else and about 10,000 live and work in Sparks. So, this is relatively balanced although if you're a transportation planner for example, you might wish that these circles were a bit more overlapping, that is people lived and worked closer to each other.

But this is the inflow-outflow so it gives you a sense of movement across the Sparks boundary, okay? So, I'm going to pause from this and go back to the slide presentation. We talked about some of the details of OnTheMap, you saw some of these things. Here are links for references to LODES and OnTheMap. Again these will be in the slide deck that you'll have available to you.

And now I'll pause for questions and answers and I'll go back to the chat and have a look. I know there were quite a lot of questions that came-in. I'll try to go through them quickly. Can I use LODES to get employment by NAICS industry at the block level?

The answer is yes but only on the work margin or residence margin. You do not have OD origin-destination access to the NAICS industry sector at the block level. And when I say NAICS industry, that sector, that's two digit for those of you who are somewhat familiar with that term.

So there's a question about Massachusetts, Alaska and South Dakota. Massachusetts has signed on in that one slide I pointed out so we do have data for all states from 2011 to 2015 and in the future we may be able to fill-in Alaska and South Dakota going forward.

Right now I don't expect us to be filling-in any more states before 2011. That's simply a matter of historical data availability. When will the federal worker data on 2016 become available? Unfortunately I don't have an answer to that question. We are waiting for the legal agreement to be approved so we can get access to that data.

Let's see, in LEHD is it possible to see the size of the firm at the work location? Specifically would it be possible to see where employees with small businesses live related to workplace as opposed to large businesses? So, we do have firm age and firm size variables. So, a couple of facets to this question.

The firm age and firm size variables first are only available on the work margin so they are not available on the origin-destination dataset so it would not be possible to see where workers who are employed at small firms live.

I will also mention that the definitions for firm age and firm size are a little different than you may assume them to be. Most of our data comes to us at the state level and is within sort of a state scope. Firm age and firm size are nationally defined so if you recently got a Walmart built in your community, the firm age of that Walmart would be the age of Walmart as a national company, not the age of that particular establishment.

We do not have establishment age or the specific site age. The same is true of the size so we have the national size. Okay, so what is the best way to work around headquartering for total employment data? I work with small geographies, typically downtowns and center cities that sometimes have overrepresented figures in specific industries.

What is the most effective way to work around this? So, headquartering is another term we have used to describe cases where firms do not report all their work sites. They may only report their headquarters and so because we don't have any other information we allocate all their workers to their headquarters when they should be distributed at their work sites.

I have seen a number of ways to do this. All of them require local data and local knowledge. That is the best way to combat this issue. In one case I saw someone who had access to data on office square footage and they identified outliers from that as in there's no way 10,000 people can be working at this place with 1000 square foot of office space.

That's just not possible and so they identified the outliers and then if I recall correctly, they scaled them down, they rescaled them. That was for their particular purpose. That may not work for you but that's one example. In general, local data, local knowledge are really the best ways to combat this issue, external datasets that you can marry-up to LODES and identify outliers and then appropriately deal with them.

Are the jobs shown-in OnTheMap in the Sparks demo the block level, is the job count available tabularly with spatial relationships? So, all the circles that you see in OnTheMap are block representations so they are block totals. The raw LODES data is at the block level so all of the residence, workplace margins and the origin-destination data are all enumerated by census blocks.

Can you select multiple places for analysis? Yes you can. I will go back here quickly to this slide. There are tutorials here if you go to this tutorials page which you can get through the Help and Documentation link in OnTheMap. There is one about geographic selection. I strongly encourage all of you to

have a look at that because there is a great deal of power to select all kinds of arbitrary areas in OnTheMap that I just did not have time to go through today.

But we do have a tutorial that will help you understand how to do all different kinds of selection. If we're okay, I know we're at 2:30 but I'll stick on here, answer a couple of more questions and what I'll say is that if I didn't answer your question, I think what we're going to try and do is get a printout of all these questions and then try and provide some kind of answer that will be made available along with the slides and other materials.

So if I didn't get to your question, I'll make sure we get to it afterwards. Any plans to integrate ACS commuting information into OnTheMap? That's a great question. Right now there are no plans but that is a question we think about quite a lot so I'll say there are no plans right now but maybe in the future.

Just to confirm exclusions 2016 to 2017 federal workers affects OnTheMap only – not all LEHD data. LODES in OnTheMap is the only dataset from LEHD that includes federal workers right now. The other datasets do not include federal workers so those other datasets are not affected by that issue.

If your region has a large number of college students residing out of state and working part-time, will it show in the data as a large population living beyond 50 miles and working for low wages? Maybe. So, college students are a good question. It depends how the data represent them.

College students quite often work in jobs that are not unemployment insurance wage covered, you know, they're like temporary workers and they can't go claim unemployment insurance but if they are employed in an

unemployment insurance sense, then we'll see the job and then it depends how we capture their residence.

We get residence information through other administrative records from other federal data sources. One big source of that data are tax records so for example if a college student still shows-up on their parents' tax forms, we'll probably identify them at their parents' home location as opposed to their campus dormitory. But again it depends on the specific data.

Is the data available for Puerto Rico? No, not right now. Puerto Rico has been a partner with us in the past and we have explored producing data for Puerto Rico. There are some data quality questions we are still trying to address related to that both spatial data quality and other administrative issues but we will continue to work on that issue.

Do you have any plans to expand the earnings category? The current maximum range really limits the usefulness of that data point. I'm going to make this the last question because this is a good one and I do want to give it a meaningful answer. The earnings categories that we have in OnTheMap and in LODES were defined 15 years ago or so.

And they were defined in a way that was meant to support analysis of low-income workers or low-earnings workers. In the 15 years since, inflation has eaten away at those categories and I agree, now they're relatively difficult to understand a good chunk of the working population.

We have been thinking about redefining the earnings categories for almost 15 years and it is still an open question but one we'd like to answer sooner than later. I would encourage you if you have ideas about what kinds of earning categories would be useful for your specific case, please send those to us.

Feel free to e-mail me and the more you explain your case, the better I'll understand it because we are trying to figure-out what we might do for the future with those earnings categories. So, I'm going to stop there. I don't know if I need to turn it over to Earlene at the end of this but thank you for listening and staying on the call for a couple extra minutes and if we didn't answer your question, we'll get to it in writing. Thank you again.

Earlene Dowell: So I would like to thank Matt for his great presentation. Join us next month on November 20 when (Heath Hayward) presents updates to another LEHD dataset and data tool, Job to Job Flows, and Job to Job Flows Explorer. Thank you so much and we hope you enjoy the rest of your day.