2022 Darwin Day Lecture

[Nina Jablonski] Greetings everyone. It's an enormous pleasure to welcome you to the 2022 Darwin Day lecture, co-sponsored by the Center for Human Evolution and Diversity and the Huck Institutes. I'm Nina Jablonski and on behalf of all of the affiliated faculty and grantees of the Center for Human Evolution and Diversity. And including my a co director, Mark Shriver, and our administrative support, Tess Wilson. We're thrilled to have you with us today. Most of you would have seen that this wonderful lecture is co-sponsored by our center and the Huck Institutes. Most everybody at Penn State knows about the Huck Institutes already, but you don't know about our center. The Center for Human Evolution and Diversity is a small but punchy transdisciplinary center that includes a variety of biologists, anthropologists, political scientists, and educators who are interested in a comprehensive view of modern humans and understanding human evolution and human physical and cultural diversity from Integrative Perspectives. We award seed grants and since last year, We have our own podcast series called Tracking Traits. Please listen. Today we have the enormous pleasure of welcoming our first virtual CED lecture for Darwin Day. Dr. Nita Bharti Nita is no stranger to many of you. She's a highly accomplished young biologist with roots in anthropology. Need a got her undergraduate degree at the University of Michigan, a combined degree in anthropology and zoology. She came to Penn State and got her master's in biological anthropology. And then moved on to get her PhD in biology from Penn State. From there She has had really a storied history. She went on to Princeton to begin her post-doctoral work and then won a prestigious Branko Weiss fellowship that allowed her to continue to develop her postdoctoral work at Stanford. Since 2016. She's been at Penn State, a wonderful coup for our institution, where she is an Assistant Professor of Biology as, as well as a member of the Center for Infectious Disease Dynamics. She's also an alloyed Huck, early career professor. So it's wonderful that, Nita is with us, but it's, it's important for us to say that Nita isn't sort of a typical biologist. When the COVID pandemic came upon all of us two years ago, Nita was one of the several people at Penn State who really launched into the opportunities for research. that the COVID pandemic offered and much of what, or at least some of what she'll be talking about today relates to her work on the COVID pandemic. It's a much richer presentation however. But among Nita's many talents as her, her ability to connect with regular people through podcasts, public lectures, and various public presentations. She has connected with a lot of people in the last two years explaining lots of stuff that's difficult to explain. So she's a real exemplar of science leadership, excellent science and science communication. Without further ado, I'm delighted to introduce Nita Bharti who's going to talk about adaptation for survival. Humans and their pathogens. Nita, take it away.

[Nita Bharti] Thank you so much, Nina for that wonderful introduction. And thank you, Mark and Nina and Tess not only for the opportunity to give this talk, but for the experience. Today, we had some nice conversations leading up to this. And so I'm excited to give this talk and be here, here, here. And one of the things that we talked about leading up to this is the idea that while Darwin is known for his work on evolution, he's also known. for his process of science. And Mark talked, Mark and I talked a little bit about his meticulous and systematic observations of the natural world. And I really liked that. So we're going to take that thread and run with it a little bit today. I'm going to provide a glimpse into the research that we do in my lab with some insight into our process. And after that, we'll return to the topic of Darwin for a short discussion on his process and placing that into a modern context. And so as I talk about adaptation for survival and humans and their pathogens, I'll specifically be talking a little bit

about adapting infectious disease management to improve health equity. Achieving health equity is baked into more than one of the UN sustainable development goals. So this is an important problem. But a lot of people are trying to address right now. And we know that health inequities persist for a lot of reasons. And just a few of them are inertia biases and data biases and solutions. Or an incomplete understanding of the interactions that are actually causing the inequities. So let's think about those interactions for a second. When we think about infectious diseases and interactions, a lot of people will think rightly about the interactions between the host and the pathogen. But there's another element here that's really important, that's equally important, and it's the environments. So this triad of host, pathogen and environment is often used to explain the fundamental interactions in disease ecology. And the environment in this triangle can include anything from climate, weather, and landscape, to political, social, and economic factors. I'm going to zoom in on the interactions between human movement, pathogen transmission and the environment and how they can help us understand a lot about behavior, disease, and adaptations. Historically, if we look back at some pathogens that have infected humans and been with us for a long time. We can trace the expansion of these pathogens across the globe with the movement of humans. Malaria, which is shown here in yellow, is widespread and it's been with humans for millions of years. And we have early evidence of malaria from infected mosquitoes found in the Dominican Republic. Leprosy is shown here in red, has been with us for tens of thousands of years. And we suspect that expanded with humans moving along these red arrows. And finally, smallpox, shown in blue, began its journey with us only thousands of years ago. And it is the only human disease to have ever been eradicated or effectively removed from the planet. And this happened in the late 1970s with aggressive vaccination strategies. This is the only human disease that most people no longer have to worry about. So when we think about movement and infectious diseases, a lot of people will immediately think of the movement of infectious or infected individuals and how they can introduce a new pathogen to a new environment or a population. And that's what we just looked at on the previous slide. It's also what we're often looking for with retrospective contact tracing. The type of movement is really important when it connects places that have very different levels of prevalence or transmission of that pathogen. And to highlight this point, the CDC shares this message on their website urging people to get vaccinated for measles, even if they never see measles cases where they live, because you can move it with you and you travel. And beyond that, we can also consider the movement of non-infectious individuals into and out of populations. This is a little bit less well studied in the context of infectious diseases, but it has a huge impact on disease dynamics and control. These kinds of movements cause changes in total population size, which is a critical factor in determining outbreak response strategies. But quantifying human population sizes and their dynamics, it's actually really hard to do. We never know how many people are in a given place at any time? But that's something that's really important to know. So if we're going to put infectious disease information in the context of outbreak management or response. We kind of need to know something about the population size and the population factors so we can assess risk, the scale of the response, and the type of the response, understanding what our options are and what's possible. For example, it's pretty easy to calculate the relative size that you would want a vaccination campaign to be around a target vaccine coverage, right? You would just calculate the number of doses you would need to distribute in a total population. But if you don't know the total population size or if that population size is changing, then things become more complicated and targets can fluctuate. Now if we think about conventional population estimates, they don't usually include details on migration or movement, because they usually provide point estimates of the number of residents in a settlement for a specific year or a period of time. If the population size is

dynamic, we have an additional source of uncertainty that can allow vaccine coverage to fall below the target level. The denominator is also important when considering health care capacity, which we usually consider a fixed maximum. How many hospital beds, how many healthcare workers, things like that. But populations move and change and adapt. They're not static. They adapt to their environments, they adapt to their pathogens, and they adapt to their own changes. So disease management strategies have to adapt as well. But my favorite examples of this is for polio management. In India. In 2009, polio was endemic in just four countries. But half the cases that were found worldwide were found in India. Due to a large population size and high birth rates, India calculated that they needed to vaccinate 170 million children per year to eliminate polio. That's a massive undertaking. So they started with the easiest and most obvious solutions and then work their way to what was harder. First, the government teamed up with international organizations like the WHO, Unicef, the Gates Foundation, Rotary Club. And they did that to secure the resources that they would need to accomplish this very large goal. And they put out broad campaigns to increase public awareness on the importance of vaccination and they used recognizable, influential faces. So if you know who this is, you know, and that all went well, then they really zoomed in on the pockets of infection. So they focused on where they were seeing the most cases. And they noticed that migrants were the highest risk group of infection. Mobile populations who were moving between places, these groups were being missed by routine immunizations. And they were also missing catch up campaigns. And that's because they weren't specifically targeted. They were difficult to access. So once this was identified, the polio campaign made a really large targeted effort to communicate with these groups, build their trust and become integrated in their networks. They also got regular people involved and they made them feel like part of the process, part of the process of eliminating polio. They were able to use that trust and that process to crowd source, the arrival of migrant groups for polio vaccination. And people weren't snitching, they were helping public health. There were no legal ramifications for these movements. And so in doing this, they were able to build spatial and temporal patterns of migrant groups. And they mapped those and kept them updated. And they did this for over three years. And now migrant populations are included in current and future routine and catch up immunization campaigns. In February of 2012, India was declared polio free. That was a huge step forward, not only for public health in India but for global health. And now 10 years later, polio remains endemic and only two nations worldwide. Some people think it's been eradicated. It has not been. Around the same time we were seeing persistent measles outbreaks in the West African nation of Niger. And when we're seeing this despite ongoing robust vaccination efforts in the cities of Niger, measles epidemics occurred every year, but only during the dry season. And the reason is largely rooted in economics and culture it would have been hard for my group to figure out exactly what was going on. Because were viewing these outbreaks through our own biases of economics and culture, most measles theory, models and solutions that we'd studied are developed around assumptions of sedentary populations, low birth rates, and nuclear household contact rates. And those factors kind of limited our understanding of the force of infection and the average age of infection and what we thought to expect didn't match what we saw in nature at all. But of course, we were working with the National Ministry of Health and they made it clear that seasonal migration for agriculture was extremely common in the area, and it occurred at very large scales. And it occurred at scales that were large enough to drive these outbreaks. So that means that a lot of people were living in low density agricultural areas during the rainy season and then they were moving to high density urban areas during the dry season. And that labor migration was causing fluctuations in population sizes. Because of that, we were seeing increases in population density and contact rates. And those were leading to seasonal

increases in measles transmission. So we kind of understood that qualitatively, but we needed to measure it. And so to do that, we quantified satellite derived anthropogenic brightness and the subtext electrification and fires at night. This is a direct quantitative indicator of changing human presence or over space and time. So you've probably seen composite images like this one. And these are used to measure population stability or population changes over long time periods, like years or decades. But since we wanted to look at seasonal changes within a year, we developed a way to look at short-term changes in populations using non-composited, serial satellite images of radiance that are captured daily so that we could measure seasonal fluctuations. And that method allowed us to detect rural, urban migration and the population fluctuations that were driving these changes in contact rates. That explained why we were seeing the outbreaks during dry season in urban areas. But it didn't quite explain what was going on with vaccine coverage. Well, it turned out what happened was that we had identified where and when to find people who were being missed by immunizations. As population sizes fluctuated. Immunization targets were not met. And so importantly, population sizes in urban areas during the dry season we're generally underestimated and too few doses of vaccine were being distributed in these populations to reach target levels of population immunity that would eliminate measles transmission. Errors in estimating the achieved coverage of vaccination campaigns end up leaving populations far more vulnerable to outbreaks, then they're prepared to deal with you think you have higher immunization coverage than you do so you're not really bracing for large outbreak. Immunization efforts in public health campaigns often have to rely on inaccurate and static estimates of population sizes. And as a result, they have inaccurate estimates of vaccine coverage levels. So improving these measures of baseline populations and measures of movement make it possible to not only manage vaccine-preventable infectious diseases and multiple infectious diseases, but also to strategically assess and strengthen health systems and capacity. And there's a lot of applications for this kind of work. We've used it for a number of things, but since we developed it ten years ago, these have been used to track human populations for a number of humanitarian reasons. These data are great but context matters, right? So satellite imagery of nighttime radiance isn't going to track all populations and especially very small populations. So we looked at places. Obviously we want to know where this doesn't work, but we were trying to solve a problem in a place where we needed a new design. We wanted to understand population, excuse me, population distribution and movement patterns in a northern Libyan deserts. And specifically, we wanted to measure access to health care and access to how pathogens we're moving through these populations that were very small and low density. And in these areas, we still see measles outbreaks. We still see malaria, and we see yellow fever. So if we look at this a little more closely, what you're looking at in the upper right is a map of population density of Namibia. And it looks kinda polka dotted because the dark areas are showing you where there's pockets of high population density. Most of the 2.1 million residents of Namibia live in urban areas. And that's what those stocks, but this dark spots are. As a middle-income country and for the most part, health care is really very affordable. If you were to be in a city in Namibia and you looked around it, the physical infrastructure, most of you would find it very familiar. Paved roads, electricity, running, water, a very urban setting. They would they would look like any other city each year. But if you were up in the Northwest corner of Namibia, things would look a little bit different. This is a very, this is an area with very low population density. And it's under served by government resources for physical infrastructure as well as health services. This area largely is occupied by nomadic pastoralists of the Himba tribe. And they have limited access to health care. Clinics are sparse and they're understaffed throughout this area. But mobile immunization services do travels through here periodically. So they arrive. They immunized

children that are present against childhood diseases like polio, measles, whatever the current campaign is. And then they move on to the next town, but they don't stay. And since they don't stay and the target populations are highly mobile, a lot of kids are missed repeatedly by these immunization campaigns. And that explains why we're still seeing measles outbreaks in these populations, despite these vaccination efforts. And of malaria and yellow fever again, are also causing outbreaks in this area. And we think that that's due to movement across this region and across these borders. So this is one of the towns in the region. You can see housing, you can see some structures and buildings and different kinds of fences and the expenses indicate land ownership. We can also see some paths or tracks in this footage and that gives us an idea put together. This gives us an idea where people live and how they might get from point A to point B. This helps us estimate how many people live in these towns at different times of year. And that's really useful for the vaccination efforts for instance knowing how many doses you would ideally want to administer in each of these towns. Or for example, estimating the size of the susceptible population during an outbreak of malaria and understanding how you might want to manage it. And we also see some divisions of labor here. So the men herd cattle and goats and they travel with their animals for large portions of the year. And the women may herd goats, but they tend to travel less than men and they conduct All of the childcare and all of the subsistence farming, which provides a lot of the staple foods. Now, we're seeing a new widespread reliance on mobility traces for mobile phones in disease research. And these data are being used to make public health decisions, including in northern Namibia and particularly to reduce malaria importations and transmission. But if these data are being used to understand populations and their movements in areas where phone data are not representative of the people and the places being targeted for health improvements, then they can actually end up not being helpful or doing harm. If biases in data are not measured or adequately addressed, we can end up measuring something other than what we think we're measuring. So in Namibia, MTC is the phone company with 90% of market share. And this is their map on top of where they They estimate that their network coverage occurs and a map of mobility traces below it. And you can see the area that, that is in question here is, is pretty sparse in both of these maps, but this is a national level network and it is the one that is being used to make public health decisions. Phones communicate with their nearest towers and you can look at phones and you can follow their movements to create these mobility traces. So we wanted to know specifically how representative phone data would be for these populations. We knew that usage was low. But it makes a big difference in determining whether usage is low or whether usage is biased. If usage is low, you can scale it up. But I've usage is biased, it can be misleading. So when we're using mobile phone data or any movement data to understand pathogen transmission, it's a useful approach because transmission events are a subset of movement. And transmission is not an observable process, but movement is. So movement is a decent proxy for potential transmission events if your movement data are representative of the population you're measuring. So we assess movements with surveys and try and assess it with phones here. And we're also assessing contacts a little more closely in this population that I'm not going to talk about that today. So we set up some data collection in this area, some temporary stations to conduct long-form interviews in these towns to really get a sense of mobile phone usage, movement and access to health care. These are our collaborators in the lower left picture, Jon, Chukrama and Jusi Mtundu. And they are Namibians who speak the language and they're really integral collaborators for this project. So we asked people about movement, where they travel, how often they travel, where they've gone, and why, we ask about their access to healthcare, how frequently they can access it or do you access it, how long it takes to get there, things like that. And then we ask them about their phone usage. Have they ever used

a phone? Do they own a phone, things like this. And then as I showed you on the map before, service, phone service isn't ubiquitous here, so there's a few places where people get phone reception. We know that introduces spatial biases in these foreign data. We're trying to get at some of that to pair it with the phone data aren't telling us from that perspective. mobile phone traces importantly, are not the same as human mobility, right? Phone data provide information on a phone that's being used in a location when reception is available while the user is completing a billable event. So there are several subsets here of what you're actually seeing. Even in locations where researchers only have a limited understanding of the representation in these data, phone derived movement patterns are often considered sufficient and representative for measuring movement when making public health decisions. The problem with this is the reasons for underrepresentation in new data screens are often due to inequities. And so the biases that lead to inequities in phone usage data can also be the causes of inequities in health that we're trying to solve. So in this case, biases inherent in these data can end up excluding the target populations from the important opportunities for public health improvements that this whole exercise is designed to improve. What we found is that in northern Namibia, only about 30 percent of the population uses phones. And that 30 percent is strongly skewed towards males. Women use phones far less than men. And we also found that people who travel, traveled in more places. Use phones more than people who don't travel or travel less. And that's what we're looking at in the plot in the lower left. Graduate student Alex Lake in my lab went further into this and he looked at phone ownership. and he did a really careful statistical analysis and I'm not going to do justice here. I'm only showing you a little bit of it. When we look at phone ownership. We see similar patterns. What you're looking at here is value of the standardized difference and that indicates how imbalanced the covariate is. So if it's more frequent in mobile phone users or mobile phone owners, then it's on the right and it's in red. And if it's less frequent among mobile phone owners, then it's in blue and it's on the left. And I'm showing you three covariates with a statistically significant relationship. The phone ownership is much more frequent among men and among people who travel to a large number of destinations. And phone ownership is low among people who reported very long travel times to the nearest health care center. Now we're getting an understanding of exactly where the biases in these data are that might be important for informing health care or public health efforts. So rather than saying that these data are totally inadequate because of bias, now that we have a sense of the bias, we can start to correct for the non-uniform bias Because these data have a lot of value, but only if we can interpret them correctly and free of bias. It's also important here that we know why people are moving and what the underlying mechanisms driving those movements are. So men are herding livestock and they're following the rains, the, the rains so that animals can graze. Years with rainfall deficits or prolonged droughts produce different movements, patterns. Women move for different reasons and remember they also do all the childcare. So if we're missing women in our phone representation, then we're also missing children, and that has important implications for childhood diseases and vaccination campaigns that we fail to acknowledge what we're missing in these data. We could end up magnifying the health disparities we're trying to reduce. But we know what those biases look like. We can do better and get around that. And you might be wondering why these locations are important for public health, since they don't seem densely populated or very connected to other locations. Well, the public health value of this data should become abundantly clear. When we look at the last mile of management or diseases like polio and measles. Or when we look at locations where diseases have emerged, what I mean is when pathogens crossover from wildlife into humans or livestock, these events often occur at the interface of human settlements and changing landscapes. The West African Ebola outbreak of 2014 illustrates the impact

that these events can have on the global health landscape. That outbreak started in a remote area, shown here, with limited phone reception, limited access to health care. and at the border of three countries. Movement patterns were really important in the spread of the infection, but are really hard to track and the phone data proved to be minimally useful. And that was true particularly in the early stages of the outbreak and when coordinate response efforts across national borders. I could probably show you a similar image for HIV emergence and several other infectious diseases. But essentially, healthcare doesn't serve global health if it's only accessible to 50 or 90% of a population, you have to reach everyone. We have to make it equitable. So while it's great that we spent a lot of time, while it's great that we spent a lot of time in detailed serial measurements and the specific area of Namibia. The big picture goal is to identify these critical gaps and biases and data and the number of locations. We want to be able to scale up our process and our findings to understand and interpret data biases in additional areas. We can do this easily in areas in Namibia were the terrain and the culture and the economic factors are extremely similar to what we've already found. But the real challenge is scaling this up to other populations to reduce health inequities in those areas. Other places that are at risk for endemic or emerging pathogens. What we're looking at here is a map of the global distribution of zoonoses that are caused by the four major pathogen types. And we know that there are some hotspots for these. And so what we really want to do is develop a process where we strike a balance between the expediency of remote methods and remote metrics and the potential biases introduced by not accounting for local behavioral factors. And so that's kind of that's kind of the blueprint that we want to follow for their, for our process as we expand this. And speaking of emerging pathogens. Regarding the risks here, in 2018, the WHO released the first list, of blueprint priority diseases, which they felt presented serious international health risks for which we, the world, were unprepared. And on this, you'll see coronaviruses and you'll see disease X. Disease X was meant to represent an unknown disease that posed a risk for research that was under-researched and under prepared for public health interventions. So when sars CoV-2 emerged, probably some people weren't that surprised. But whenever diseases emerge, the first line of prevention is always behavioural. With emerging pathogens, behavioral interventions are really important because pharmaceutical interventions don't exist yet or not available for novel pathogens and they haven't been developed. So what we often do is limit movement, reduce contacts, and target personal hygiene. So what you're looking at here is commercial air, airline routes, train station, foot traffic, and obviously personal hygiene, handwashing is a very frequent target for behavioral interventions. And we do this because frankly it works. Behavioral interventions are really effective. Prior to 2014, behavioral interventions were the sole tool for managing Ebola outbreaks. And they kept all Ebola outbreaks under 500 cases from the first known outbreak, in 1976 to the outbreak in 2014 that prompted the development of the first Ebola vaccine. It's hard to measure the uptake and impact of behavioral interventions, and that's a problem. So there are top-down suggestions that are a little easier to track but are short-lived. And then there are individual adoptions of behavioral changes and their persistence and those are really hard to identify and measure. And after living through this pandemic for two years, people will be really familiar with both of these types of interventions. So we tried to look at these really close to home. We're looking at pandemic behavioral changes in Central Pennsylvania and what we can understand about their impact on disease transmission. We want to know how compliant people were with top-down restrictions. We want to know what they changed about their individual behaviors and how long those changes lasted. And we want to know how much it impacted disease transmission. So we know qualitatively that things have changed, right? The last time that we had a Darwin Day lecture, we were all in attendance in

person. We were unmasked. Visitor traveled, the speaker traveled to visit us. And it was very different from today. So we know things have changed qualitatively, but we really want to measure these quantitatively. So how do we measure movement and behavior? In central Pennsylvania? We started doing this in near real time and the summer of 2020. This is my former post-doc Christina Faust she has since started her own lab, in Glasgow, that we started with the case data during the earliest days of the pandemic. And we looked at how those responded to these restriction phase. As you remember, we had red, yellow, and green restriction phases that were increasingly lenient, that involved business closures, stay at home measures, closures of schools and daycares and things like that. And then those were lifted from red to yellow to green. So we looked at the case data and central Pennsylvania through these restriction periods. And when we thought about how to measure movement in this area. A programmer in my lab, Brian Lambert had a really good idea. He knew about these traffic cameras that are managed by PennDOT and a couple of other sources. The township manage one, manages one, and the university manages one and there's a few others, but There's about 22 traffic cameras that surveille the roads in Centre County. And he knew about these because he admitted that he would use them on football weekends to see if it was safe to go out or not. And this is genius. It's a trick that I totally stolen. But on top of that, it proves to be a really valuable tool for how we might be able to measure movement in Centre County. There are, as I mentioned, over 22 cameras that look down on the roads, in Centre County. And their spatial placement is shown in the map in the blue and yellow triangles in this center plot in the upper row. The blue triangles are showing us connector roads like highways, freeways. And the yellow triangles show us internal roads or surface traffic, surface road traffic. And so Brian's able to do is he's able to grab an image for each camera every 20 seconds from these public video streams. And he can basically quantify the number of cars and pedestrians, or vehicles and pedestrians in each image. And he set up a pipeline to completely automate this process. So we're able to look at how many cars are on the road every 20 seconds. For all of these cameras throughout the county. When we looked at these data during the first summer of, of COVID, what we found was going from the red phase when we started this data collection to the yellow to the green at the traffic increased. And that's what you're looking at in the plot in the lower left. So people are generally compliant. They stayed home and they were told to stay home. They went out a little bit more as they were permitted to do so. And then when we looked between the red and the green phases across different types of roads, we notice that traffic just kind of increase across all of them, probably a little bit more across internal roads, but people really were showing that they were compliant with these measures. And when Christina looked at these volumes of a vehicles against the COVID cases, she found that there was about a 10 day lag between traffic volume and increasing and an increase in COVID cases locally. And while testing was really low during this period, we would expect, but severe cases would reveal would reveal some of that testing if we were way below if the data were way below the reality and we had really low hospital occupancy. So we think that it's probably the reported cases are probably a decent proxy for the actual cases. So we saw this and then we kept monitoring, monitoring traffic and cases. And we looked at this through the student return. And this is where things got interesting. So the traffic data and the top row here didn't actually detect a huge peak of student return. But we knew that at least 30000 students return for the fall semester of 22 of 2020. Sorry. But we saw a huge increase in the number of cases about two weeks after students returned. Based on that, we were able to calculate that what we probably had experienced was a large influx of susceptible individuals and a very small influx of infected individuals. And then a lot of transmission. When we disaggregated these early cases, it became clear that this outbreak wasn't affecting everybody equally in Centre County. Specifically, what we're looking at here

on the left, cumulative COVID cases across all counties in central Pennsylvania. So the seven central counties in Pennsylvania, and those are shown in the gray and the black. And then in purple we see all the Centre County, blue, we see the students in Centre County. And in red we see the non students in Centre County. And what we see immediately is that the students cases in blue increased really rapidly and increase to a very, to be a very large pandemic. And the non-student cases stayed pretty low. They increased at the same trajectory as all the surrounding counties. So they didn't experience this large increase in cases that the students experienced right away. There's something here that kind of surprises and we couldn't get our finger on it right away, but we did later and I'll come to that in a minute. When we break down the total per capita cases in each of these areas, what we find is that Centre County as an aggregate, nothing really jumps out. That's the purple point and the purple line on the left. So if we look at the aggregate of cases Centre County, looks just like another county. But if we break it down into students, a non students it's clear that the students have a much higher per capita rate of cases from the non students. And the, the points that you're seeing in red indicates some uncertainty around the a number of cases in the non-student population. But even at its highest estimates, there's a lot fewer cases per capita in non students in Centre county than there are in other counties and other neighboring counties. So we use a dig into this a little bit. And one of the things that we needed to do was get another estimate on movement. Around this time, the company safegraph started to make mobile phone usage data available to researchers to study the impacts of the pandemic across the US and Canada. So we started looking at a different kind of movement with these data. And particularly these data enabled us to look at measures of movement before the pandemic had ever started. And that's what you're seeing on the far left and top graph. We can't do that with the traffic data because we didn't start collecting it until the red phase. So we can measure relative changes. Well, we don't quite know what the pre-pandemic baseline would've looked like. So this helps us calibrate some of that. What we're looking at in this top plot is the rolling 70 mean of total daily visits. The points of interests. That could be anything from retail locations to parks, to campus buildings, points of interest, points of interest per county. And the areas that are shaded in gray correspond to times when students are in Centre County. And they correspond to the times that are shown on the plot below, which I'll explain in a second. But we're able to see here with the phone data is of course, a pre-pandemic baseline of what visits to points of interests look like. But we're also able to see that the phone data appear to be catching, capturing the return of students for the fall of 2020 a little bit better than the traffic data. What's interesting is that the phone data capture a smaller percentage of the population when the traffic data. So what we think we have is a really nice set of two complimentary datasets. The phone data that capture the students fairly well, and the traffic data that capture of a non students fairly well. And we've done some more analyses into that. And it really does look like that's kinda what we're seeing. So that's really useful. So I also look at visits to points of interests throughout their pre pandemic and the pandemic times with these data are also able to look at how much time people spend outside of their homes. And we can break this down by county and we can look at non students and county totals. And what we see is that on the far left set of bars in this plot, on the bottom, we're seeing pre-pandemic measures of how much time he will spend outside of their homes. And in purple we see all of Center County. And in red we see the non students of Centre County. All the other colors are showing us the other counties around Centre County. So the first thing we notice is that when pre-pandemic times people in Centre county, spent a lot more time outside of their house than people in other counties. But then it very quickly becomes apparent that during the pandemic, People in Centre County spent a lot more time at home than they do than the people who live in other counties. So

there's something important here and we get into it. We did an analysis on what may be causing these differences. Will also see that these differences look like they line up with per capita cases during this time period of COVID. and essentially we're able to link some of these elements to median household income. So if you make more money you might be able to work from home. Pretty much just a, just a way of measuring pandemic privilege. So we were able to figure that out but that brought us back to our confusion about the issue of why the COVID cases are so asynchronous between students and non students in Centre County. I'll bring you back to that plot really quick. When we're looking at, is again, an outbreak that looks like it's pretty much decoupled between students and blue and non students in red. We realize the reason that we were, we were caught off guard by that pattern is because if we look at influenza on the right, we see seasonal peaks of influenza. They're highly synchronized between students and non students. The student data are shown in red and in the red points. And the state wide data from Pennsylvania are shown in these black lines. And so there, we're seeing influenza, another respiratory virus. We're seeing those outbreaks at the same time. But when we saw COVID, we didn't see curve at the same time between the two populations. Interestingly, the year that we had this giant COVID outbreak, our first massive COVID year or that would have aligned with influenza, we had no influenza. And that's essentially because the COVID behavioral interventions were so effective at limiting influenza transmission that we just didn't have an influenza season. But we had COVID and we had COVID that didn't line up between these two populations. So we've received some NSF funding looking forward to understand the impact of behavioral interventions from COVID and how they are affecting other respiratory viruses including influenza. And this is Dr. Nikparvar a new person in my lab who's going to be examining these. And he's got a bunch of additional data sets including self-reported behaviors and longitudinal surveys. And he's going to be looking at mixing, movement and contacts within and between these populations and tracking COVID and influenza. Moving forward. We think that there's a lot of inequities that are driving these patterns. Already mentioned that we have identified some economic reasons that we see in differences in the ability to stay home or avoid COVID. And then we've also seen some interesting patterns with crowding and density. And we think that that's really kind of what made this so problematic for the students that they basically didn't have the resources to quarantine or isolate from, wouldn't really set them up to succeed that way. So we're looking into this and the idea is to be able to come up with lower cost, lower impact, targeted behavioral interventions. To reduce other respiratory viruses in the future. We're really exciting avenue for research. So a quick summary of the projects that I just talked about. We can improve health equity and we really should. We have to know what our data don't show and we have to measure bias and not ignore it. And that includes our own biases, biases in our data. And the biases that go into the methods that have been developed to address infectious disease management. On that last points, let's return briefly to Darwin's scientific process. Darwin's famous text On the Origin of Species is, I think his most famous book, it was the one where he described the early roots of evolution in animals and plants. And this was before he or anyone else knew anything about genes. He drew remarkable conclusions from careful observations and scientific processes. And he really embraced that science as a process of constant improvement and correction as, as a scholar and a scientist himself. So he revised the Origin of Species nine times after it was first published. And in honoring the true spirit of science, and the scientific process, we're going to spend a few minutes talking about what Darwin got right and what he got wrong. And we're not the only people who are going to do this. This has been a conversation that has, that, that is ongoing and that is really important to scientists that we have. It's our job to always be revising and improving our working hypothesis. So it's worth noting where Darwin fell short when he seemingly abandoned too

much of his own scientific rigor in the interpretation of human evolution. And specifically with regards to sex and race. And he did this in his two-volume publication on the Descent of Man and and Selection in Relation to Sex. And in this book, in these books, these two volumes, he really kind of tackled human evolution. What, you know, what he thinks happens here. He didn't fully address this in The Origin of Species. So this is where he gets into it. The book cover on the right is a book that I highly recommend. If you find this topic interesting. It really breaks out in detail a lot of where Darwin's sort of abandoned his scientific process throughout his, throughout his writings and this book. We'll go through just a couple of them that I find really interesting. So Darwin, on the difference between men and women, the difference between the sexes. He wrote in The Descent of Man, that women were intellectually inferior to men and inferior. in just about every way. He wrote that women could improve their intellectual capacity, their skill level, whatever, through sexual selection. But then he kind of talks himself around to the point that it will, if women could do that, then men would do that too, and they would just simply improve more than women. And women could never achieve equality. And he said this Despite with his own scientific observations showed and what his own personal experience experiences were, his data didn't support this his observations on men and women, didn't support large biological differences that would lead to that kind of hierarchy that he's talking about. And in his personal life, his daughter with his primary editor for a lot of his most influential scientific writings. So he didn't really seem to follow the data when he made these conclusions about how man was superior to women. But he's somehow arrived at these conclusions anyway. Similarly, Darwin observed similarity amongst races. And famously, he was strongly against slavery. And he argued that all human races were members of the same species. Let's not something that everybody of his time thought but he thought that, but he didn't believe in racial equality. He believed that there existed a racial hierarchy. Despite the fact that these quotes on the left from the Descent of Man show clearly that he is documenting similarities among races. He's documenting similarities in their biology and their physiology and even in their minds. He's having a hard time really drawing the kinds of conclusions that we would expect to follow here. So despite saying that people are so much more similar across races than they are different, he concludes that there must be some differences in intellectual or moral, or social faculties. And he frequently misinterpreted people and cultures that were unfamiliar to him as inferior, simple or primitive. Well, why did he do that? Why did he make these really careful observations? And then draw conclusions that weren't really supported by his data and his observations. His own data didn't support his conclusions on race. And you can read this in his text, The Descent of Man. His own personal experience did not support his conclusions on race his most influential and favorite teacher was a black man who taught him all about how to preserve specimens, which was absolutely integral to his ability to write The Origin of Species and to move specimens across the world and, and look at them and, and, and really compare them. He wasn't the only white male scientists during his time to present this way of thinking. He, he had some contemporaries who, who also felt this way. But he also has some contemporaries who followed the data and believed in equality between races and sexes In The Descent of Man. Darwin failed to recognize and measure his own biases of human sex and race. He incorrectly interpreted his observations through his own personal biases. Somehow he was able to overcome those biases when he was interpreting the evolution of plants and animals. And he presented his theories even though they conflicted with the church. And those theories that followed the data and the evidence they still hold up today. At our efforts to honor Darwin as a complete scientists, we have to look back not only on his successes, but also where he fell short and how we can learn from his errors to do otherwise would obviously be intentionally emissive and disingenuous, also sends a dangerous message, either of willful

ignorance or intentional support. Some of his unsubstantiated findings that we can honor the findings and the Origin of Species and understand a massive impact that they continue to have today. While we recognize the errors where his process was abandoned and his conclusions fell short, The Descent of Man and the massive impact that they continue to have. So with that, I'll come back to my summary and we'll focus on the last point. Measure bias wherever it is, whether it's in yourself, whether it's in your data, wherever you're seeing bias, address it head on. And that'll make you a better scientist. I'm going to take a quick second to acknowledge collaborators and funders. Think I talked about work everybody except Kelsea Baronowski a graduate student in my lab doing amazing work on the emerging virus, Hendra virus and landscape change. So hopefully you'll have a chance to check that out. Some of our funding sources, collaborators and data sources. This is all the work that this is the funding for the Namibia project and these are all collaborators. And this is the funding for some of the work on COVID and Centre County and surrounding areas. And these are our data sources. And data for action is project that's happening at Penn State that I'm a part of where we are collecting these really valuable data from students and non students throughout the county. And with that, I thank you for listening and I think I can take questions. Indeed, you can.

[Nina Jablonski] If you have questions for Nita please put them into the Q and a, which you can toggle at the base of your screen. And Mark Shriver will be happy to read those questions. Nita, while people are marshalling their questions. I'm interested in just asking you a quick one about in your work in non-Western contexts, what kinds of reception has there been to what you're doing and what you're trying to do. We hear occasional reports about inimical reception that is given to people who are involved in vaccination campaigns or gathering information on epidemiology. And I'd be interested to know your experience and how you see these things being more successful in the future.

[Nita Bharti] Yeah, so that's a great question. And it is a question that I actually really love to answer them. Thank you for asking that. I think in any situation, if you're going to go to another community or another culture, you have to, you have to do that with the intention of understanding their priorities and solving the problems that are their priorities. If you bring your priorities into somebody else's community. Yeah, you're going to be met with maybe not the warmest reception because if there are problems that need to be solved, That's kind of where you should start. So I think understanding the priorities of the communities that you're, that you're working, that you're working with and maybe starting with problems that need to be solved. And really focusing on problems I need to be solved is not just good research practice, but it's, it's a really fundamental way. To work with communities and to make actual improvements to solve real problems.

[Nina Jablonski] Great, Thank you.

[Mark Shriver] So we have our first question, Nita and it's in honor of Darwin's day. What do you think the biggest takeaway from Darwin's scientific procedure is?

[Nita Bharti] I think that there are probably quite a few good takeaways from Darwin and has processes. I think hopefully we've seen a little bit of what to do and what not to do or how you can lose yourself astray. But I think there's, there's something really important about careful observation. So really take the time to figure out what you're seeing, record it, and process it. Think about it and, and, and synthesize across different data sets to really make sure you're getting a full picture. I think that there's, I think that there's a lot of really interesting things that Darwin did when he really followed his process

that have almost more than remarkably, very surprisingly stood the test of time. I think we can learn from him that way.

[Mark Shriver] Could you see an effect of football games, mass gatherings, and was there a difference in peak of peak in traffic and surrounding counties versus student incidents?

[Nita Bharti] So we a 100 percent can see football games in basically every measure of movement throughout this area. Yes, me see them loud and clear and actually kind of useful in the sense that stadium attendance is recorded. And so we can look at how many visits were to the stadium during the game and calibrate what proportion of game attendees were seeing in the same breath, in the phone data. So we definitely see football games and other mass gatherings so that they're moving data do a remarkably good job of tracking large events. And obviously the largest of those are football games. Was there a difference of peak in traffic and surrounding counties versus student incidents? I'm not sure I understand that question, but I'll talk about what

[Mark Shriver] I think it's asking for Were infections associated with the games through the data

[Nita Bharti] So the first year when we couldn't really cancel football, we have football without spectators. We. Yes or no. So there's a lot of other things that would've happened on the weekends that might have contributed to transmission among students. But really I think for the most part what we saw with students was just the, the effects of crowding. I don't think it really I don't really think they were armed with the tools necessary to, to quarantine or isolate. And so I just think the crowding in that and the density within housing was so overwhelming that I'm not sure the activity levels really made a difference.

[Mark Shriver] So just, just a branch off of that. Are there, are there any recommendations you might make to help protect students and future outbreaks or even just for the flu next year.

[Nita Bharti] Masks. For starters, their masks would be great. for flu and certainly for COVID as we move forward. My recommendation for this really is just get tested, often get tested early. And as soon as you have a result to act on, act on it. Our testing capacity is much better now than it was before. And I think, you know, people can fully take advantage of that. And that is a huge asset. If we think back to like early HIV messaging. That was a huge part of getting ahead of the HIV outbreak and making progress on that pandemic. And it's really just basic public health get tested and know your status.

[Mark Shriver] Okay, next question. Great talk. Thank you. I'm curious whether the lagged curves of students vs. non-student cases in Centre County could be indicative of returning students, seeding cases in the community. Additionally, were there ways to account for public transportation use in the traffic data? Can you tell whether buses were more crowded or not?

[Nita Bharti] Yes. So the first part of this question is really good question. And we looked at this very, very carefully. What we really saw was when students returned, only students got sick. It took a very long time or non students to start recording an increase in COVID cases and those outbreaks the outbreaks. And then in the non students and Centre County looked exactly like the outbreaks in neighboring counties where there were no students. So we really think that transmission between students and non students was extremely minimal if at all. And that was surprising to us because we don't see that with other respiratory viruses. So we were, we were kind of shocked by that. That has to a 100 percent be the the result of behavioral interventions. So students likely still had to interact with

each other. But non students were able to not only interact less with other non students, but also to interact less with students. And so we didn't see those cases jumping back and forth between students and non students. I don't know that that'll be the case moving forward. I think we're going to lose some of that signature, but that's part of what we're looking at. So I don't think that the data are indicative of returning students seeding cases in the community. I think we're seeing quite the opposite. That was a really nice question. I'm just so public So we've looked a little bit of public transportation. Early in the pandemic public transportation kind of was shut down and really reduced. So that ended up not really being a huge factor. You can look at the volume of or that you can look at passengers on a bus or how crowded buses with the with the public transit data. But didn't nothing really that there was really no relationship there. We looked at that kind of quickly and we moved away from it. It really did not seem to be a contributing factor. Good question though, we definitely looked into that.

[Nina Jablonski] I'll ask another question because I'm fascinated by your point about Darwin's biases. We're very quick to point fingers at old dead guys, and less willing to point fingers at ourselves. Knowing what you do, not only about your own research program, but the research that is ongoing on COVID and COVID epidemiology. Give me a time machine. Look at how we might view our own biases today. Because certainly we are bringing some biases about, about people, their behavior, their thoughts into the way that we collect data, even about seemingly objective things. And so I'd be interested in your take on what our biases might be today.

[Nita Bharti] Yeah, so this is a great question I'm going to probably need a minute to think it through? But yes, we bring with us biases to everything, including the data we collect and how we interpret those data. With COVID, I think it's a really, a really sort of low, low level answer to that question, for starters, would be the issue of kind of what we just talked about with students and non students in Centre County. We really thought one of the, one of the prevailing theories heading into the fall of 2020 across college campuses was that if you bring a bunch of undergraduate students back to campus, they're going to bring with them a lot of COVID and it's going to impact the surrounding communities in a very detrimental way. That's really not what we saw. But that is definitely a bias that drove that research, right? So we wanted to see if it was there because that's what a lot of people expected. We didn't know what to expect, but that's a bias that definitely existed. And, and I think that was an important one to look at. I think a lot of our biases early in the pandemic proved to be really dangerous. And I think some of our early biases lead to some of the really differential rates of morbidity and mortality that we saw in communities of color, certainly black communities and brown communities versus other communities versus white communities. And I think we had a lot of biases about when it was okay to absorb risk, and especially when it was okay for other people to absorb risk. So I think that that was a big problem and I think that there were also some really public statements made by politicians about lifestyles that were contributing to risk, that really weren't lifestyles at all. And they were things that would have aligned with what people had about racial assumptions. I think some of those things are really dangerous. And I think they probably kept us from correcting some of the differential rates of morbidity and mortality that maybe could have been reduced earlier.

[Mark Shriver] Do we have anymore questions? Okay, Well, we're about 15 minutes after the hour. I guess. Let me thank everybody for attending and thank you, of course, Nita for speaking to us. Very, very fascinating talk. And thank everybody for participating. Really appreciate it. I want to thank the Huck Institute and CHED too for being able to get the seminar series going.

[Nita Bharti] Thank you so much. This was wonderful

[Mark Shriver] Bye everybody.

[Nina Jablonski] Bye.