

Analysis of Project Flow and Use of RDC Services for the Project:

Examining Workplace Geo-Coding for Travel Forecasting

Using the American Community Survey

Prepared by

Deb Niemeier, Ph.D., P.E.

Prepared for

U.S. DOT Federal Highway Administration

July 2004

Table of Contents

Introduction	3
Project Tasks	4
Additional Aspects of RDC-Based Projects.....	6
Main Issues and Stumbling Blocks.....	10
General Observations	15
Appendix A: Detailed Timeline	18
Appendix B: Disclosure Forms	22

Introduction

The decennial census long form has traditionally provided supplemental demographic and socioeconomic data for one out of every six households. This information is used extensively by the transportation planning and travel demand forecasting communities. The Census Bureau has begun implementation of the American Community Survey (ACS) as a replacement for the decennial census long form. The Census Bureau currently has no plans to conduct a long form in 2010. The ACS will include 3 million housing unit addresses each year, such that over a 5 year accumulation period, 15 million housing unit addresses will be surveyed, or approximately 12.5% of addresses. The 2000 decennial census long form included approximately 16% of addresses.

In order to evaluate the suitability of the ACS for the purposes of transportation planning and forecasting, FHWA commissioned a comparison study between the 1999-2001 ACS data and the Census 2001 San Francisco County estimates to identify any inconsistencies or important differences between major workplace-related variables, particularly those used extensively in the transportation travel demand forecasting and planning processes.

A primary objective of the study was to assess the availability of sufficiently representative data over each wave of the survey. In particular, a comparison between the amount and quality of geo-coded workplace information available by population and traffic analysis zone (TAZ) geographic size/location variables over each survey pass and the census long form data was desired.

The commissioned study began in March 2002 and culminated in early summer 2004. The study was conducted through the Census Regional Data Center (RDC) at the University of California (UC), Berkeley. The purpose of this report is to outline some of the issues and difficulties associated with conducting research at the Census RDCs. Although some of the experiences discussed in later sections refer directly to the UC Berkeley RDC, the process-based experiences are likely to be similar across RDCs. The report begins with an overview of the basic tasks undertaken as part of the commissioned study. In subsequent sections, discussion of certain key work efforts are expanded upon in order to further highlight

stumbling blocks and procedural difficulties in accessing and analyzing the ACS data at the RDCs. The report concludes with a summary of observations and possible suggestions for improving analysis of ACS data in the RDCs.

Project Tasks

The commissioned project had a study scope organized into three main analysis tasks and one largely administrative task. The first three analysis tasks are discussed together because they are closely related and somewhat dependent upon one another:

- Task 1: Examine the transportation-related effects of data accumulation
- Task 2: Classify geo-coded workplaces for transportation
- Task 3: Examine the MAF updating procedures

The ACS is an accumulated dataset; that is, full data are not available until the end of the five-year sampling period. Historically, the Metropolitan Planning Organizations (MPOs) have relied extensively on the Census Transportation Planning Package data for model updates and planning exercises. Thus, the availability of the ACS data for use in travel model development is a function of the timing of the ACS survey waves relative to the MPO model updates. For the purposes of upgrading travel demand models, the way in which data accumulate in each wave of the ACS survey may have an important effect on the availability of workplace information for certain TAZs at model update times. Ideally, the number and representativeness of geo-coded workplaces should be sufficiently robust such that model updates are not wholly dependent on the ACS survey pass year. For these tasks, a subset of TAZs was to be examined using the 2000 Census, information from the Bay Area transportation model, and the ACS workplace information. Although the primary interest, in terms of spatial resolution, was at the TAZ-level, it was not clear at the start of the project that this level could be supported with the available ACS data.

Among the questions to be addressed in these tasks were: 1) how many geo-coded workplaces are available at each sampling phase for use in the travel demand model; 2) are geocoded workplaces sufficient in number, space, and quality for use in the travel demand models, and 3) what are possible methods for imputing or deriving weights for estimating workplace destinations by small zones or zones with very few, and perhaps less

than optimal representativeness geo-coded workplaces. That is, the primary interest was in the amount and quality of geo-coded workplace information available by population variables and, ideally, by TAZ geographic size/location variables over survey passes relative to the 2000 Census long form information. The products of the first two tasks were to develop quantitative assessments of the workplace information. The quantitative information would be used in Task 3 to potentially help Census to refine the sampling frame and identify geo-coding priorities for each ACS survey pass relative to transportation needs and requirements.¹

To construct the current sampling frame, the Census Bureau uses a national Master Address File (MAF), which they maintain, and was assembled by generating a composite match file using the U.S. Postal Service (USPS) Delivery Sequence File (DSF), the 1990 Census Address Control File (ACF), and the TIGER files (American Community Survey, 2002). The MAF can be automatically created for areas with city-style address mail delivery systems. (The MAF must be generated differently for non-delivery areas). The MAF serves as the sampling frame for the ACS.²

To sample households in the ACS, each month the Census Bureau plans to select a systematic, random sample of household addresses, representing the entire U.S., from the most current MAF for the ACS. A larger proportion of addresses will be sampled for small governmental units (American Indian reservations, counties, and towns). The monthly sample size is designed to approximate the sampling ratio of Census 2000, including oversampling of small governmental units. In short, the current sampling method relies mostly on households, not workplace destination, and consequently, the impact of the sampling method on workplace location for purposes of transportation forecasting is not well known.

One of the critical elements that the Census Bureau has identified as important to the overall success of the ACS is the ability to keep the Census Bureau's MAF up-to-date and accurate from year to year. The MAF serves both as the main source of the housing unit

¹ Currently Census geocodes 75% of the workplace information.

² See, for example, *Accuracy of the Data (2000)*, www.census.gov/acs/www/Downloads/ACS/Accuracy00.pdf

sample for the ACS and plays an important role in the editing, weighting, and data tabulation process. The Census Bureau has been developing a new program called the American Community Survey - Coverage Program (formerly called the Community Address Updating System), which has two major objectives: 1) to obtain address information about new housing units and add those units to the MAF; and 2) to correct and update the existing addresses in the MAF.

Based on the results of Tasks 1 and 2, the third task was aimed at developing recommended improvements to the ACS Coverage Program. Specifically, the intent was to recommend ways in which the sampling protocol and frame could be refined to improve the availability of workplace information for transportation planning and travel demand modeling.

The final study task included in the study scope was the preparation of a final report. During the course of the commissioned study, some elements of Tasks 1 and 2 were completed, but not enough to allow completion of the study. Despite the relative straightforwardness of the study design, a significant number of stumbling blocks were encountered during the course of the project. Most of these elements created serious project delays and in the end, compromised project completion.

Additional Aspects of RDC-Based Projects

In order to undertake a study in any RDC, there are a number of steps that must be incorporated into the basic project scope. That is, without completing each of these steps, the project itself can not be undertaken and/or completed using RDC resources. These additional tasks can be divided into pre- and post-analysis phases (Table 1). Each of the steps has associated with it additional time requirements that are often above those normally budgeted into data analysis studies. The times in Table 1 refer to the approximate times associated with this specific study. In the following sections, each of the tasks are briefly described along with some of the issues that arose during the process.

Briefly, acquiring approval to access RDC resources requires that the study concept initially be vetted in pre-proposal form to the RDC Director, and if accepted, in full proposal form to the CES Proposal Review Board. The expediency of this review process depends on several things. First, the pre-proposal application is geared towards researchers requesting use of the economic databases. The researcher is requested to specify a CES database (from a pull down menu) and then manually enter any other data that might be desired. Each type of data have different issues associated with using it.

When the pre-proposal application was originally submitted for this study through the web interface, it was unclear whether a CES data set also had to be specified in addition to the ACS data. The first time the pre-proposal was submitted, only ACS demographic data were requested. There were problems with the web server and the UC Berkeley RDC Director was unable to access the pre-proposal to accept it. A second pre-proposal was subsequently completed and per the RDC Director's instructions, workplace CES data were also requested under the required field. This created a new problem because each CES data has specific limitations associated with using the data, and these limitations don't necessarily apply to the ACS data. As the table in Appendix A shows, this process required several weeks to complete.

Once the pre-proposal has been accepted, the researcher is typically invited to submit a full proposal. This requires very little additional administrative time beyond that required to address any concerns identified by the RDC Director during the pre-proposal stage. The full proposal is reviewed first by the RDC Director, at which point some changes may again be requested. The proposal then goes to the CES Proposal Review Board and possibly to outside reviewers. The Board meets only a few times each year and proposals are given full approval, partial approval requiring revisions, or denied. For this particular commissioned study, this step took several months to complete. While this might seem somewhat of a long review period, even longer review times have in fact been documented.

Table 1. Breakdown of Additional Tasks Related to Use of RDC

Project Steps	Approx Time Required¹
FHWA Authorizes Study (Based on reviewed proposal)	
Pre-Proposal Submitted to CB	2 h
Full Proposal Sent for Review by RDC Director	
Pre-Proposal Accepted by RDC Director	
Full Proposal Reviewed by CES	
Full Proposal Revised Per RDC Director	1 h
Full Proposal Accepted by RDC Director	
Full Proposal Reviewed by CES Proposal Review Board	
Revisions to Full Proposal Based on CES Proposal Review Board	4 h
Full Proposal Submitted/Accepted by RDC/CES	
Special Sworn Status Forms Completed	} 6 h
BC-1759: Special Sworn Status Form	
SF-85: Questionnaire for Non-Sensitive Positions	
OF-306: Declaration for Federal Employment	
86C: Special Agreement Checks (SAC)	
Holmes / IRS Memo	
Fingerprinting (Census can not use the livescan option)	3 h
UC Berkeley Human Subjects Application and Review/Revisions	15 h
Work Contract Prepared/Accepted	3 h
Thin Client Paperwork	3 h
CES Computer Account and Data Request Paperwork	1 h
CES Researcher Handbook Review	2 h
Pre-visit to RDC (Required)	8 h
Title-13 Training Tutorial	
Training certificate faxed to DC	
DC staff receives certificate, issues ID/password.	
Researcher confirms all necessary data are available	
<i>Analysis at RDC</i>	
Disclosure Request Form ²	3 h
Benefits to the Bureau report ²	2 h
<i>Analysis Results Available from Census</i>	

¹ Administrative time above that that might normally be included in a project budget

² See Appendix A

Once approval of the full proposal has been received, the RDC Director provides the material necessary for acquiring Special Sworn Status. This material, outlined in Table 1, requires fingerprinting in addition to information related to conducting a background check.

Interestingly, one of fairly time consuming stumbling blocks for the study conducted at the

UC Berkeley RDC related to the use of livescan versus traditional fingerprinting. Nearly all of California law enforcement and governmental agencies now use livescan. With livescan fingerprints are automatically electronically filed with the appropriate agency, which means that a digital receiving address must be specified. The electronic delivery obviously speeds up the process; however, the Census does not yet accept livescan fingerprints. Thus, for places such as California, finding a location using the traditional fingerprinting process can add additional time.

Concurrent with the Special Sworn Status process, researchers must also file a human subject's application with the Committee for Protection of Human Subjects (CPHS) for permission to pursue RDC-based research at UC Berkeley. This process requires completion of a completed Financial Conflict of Interest Checklist, CPHS Application Coversheet Form, and a detailed description of the research. Acquiring IRB approval can also be somewhat more difficult if the researcher is not associated with the UC Berkeley campus. However, on the positive side, for most ACS-related studies, an IRB exemption will eventually be granted.

In addition to these steps, researchers must also complete a work contract before starting. This is basically a shortened, more succinct version of the full proposal. Although the purpose of the work contract is not exactly clear, it seems to serve as the catalyst for securing the RDC funds necessary to cover the research. In the case of this study, the RDC funds were internally transferred between the Census ACS office and the CES/RDC group. Finally, there are thin client and computer account forms and a review of RDC procedures that must be completed.

Once all the necessary approvals have been granted the researcher is asked to schedule a visit to the RDC for an orientation session. During the orientation, a self-guided interactive session and testing on Census related confidentiality and data management issues is conducted. Once this session has been completed, the RDC Director contacts Washington, DC and receives the necessary information for acquiring log-on privileges, at which time the researcher is asked to review and confirm that all requested data are available.

At the completion of the project, all data output and data runs must be submitted through RDC for disclosure. At this point, there are two forms that must be completed (Appendix B). Completing these forms is time consuming and usually requires another RDC visit to gather all the necessary information.

Main Issues and Stumbling Blocks

There are a number of stumbling blocks that arise when working on ACS data at the RDC. These can be roughly organized into proposal issues, data management and analysis issues, and paperwork issues. The approximate time lines and proportion of project time by activity for this particular study have also been documented in Tables 2 and 3.

Table 2. Timeline of Major Steps

Step	Timeline
Pre-proposal Submit/Review/Accept	Early-April 2002 to Late-April 2002
Full Proposal Submit/Review/Accept	Late-April 2002 to August 2002
Special Sworn Status	Early-August 2002 to Mid-March 2003
<i>Data Delivered to CES</i>	<i>Late-November 2003</i>
<i>Data Received at RDC</i>	<i>Mid-May 2003</i>
Human Subjects Filed/Approved	Late-August 2002 to Early-August 2003
Work Agreement Prepared/Reviewed/Accepted	Mid-March 2003 to Early-May 2003
<i>RDC Orientation Visit/Authorize to Proceed</i>	<i>Mid-June 2004</i>
Disclosure Filed/Comments Received	Late-January 2004 to mid-April 2004

Table 3. Approximate Allocation of Contract Time

Task	Allocation of Project Time (%)
Pre-Analysis Tasks (Table 1)	34%
RDC Visits	29%
SAS/ArcView Script Dev. And Testing	15%
Gen Admin (incl report prep)	7%

Proposal Issues

In most studies, researchers usually develop fairly detailed proposals for review by funding agencies. In order to use the RDC, these proposals must be extended to address a number of Census specific elements. Along with proposal information, considerable documentation must be provided on how the study will directly benefit the Census. In addition, based on the

subsequent CES Proposal Review Board comments, modifications may often be required of the original agency accepted proposal.

In the case of this study, a proposal had been accepted by FHWA at the time the study was commissioned (as will generally be the case). After the review by the Proposal Review Board, only minor changes had to be made to the original FHWA funded study in order for CES to approve the study. However, at a later stage in the project, when difficulties related to acquiring certain GIS information were encountered, CES staff indicated they had not really understood the scope or complexity of the approved analysis³. This would tend to suggest that study proposals can be approved and still run into difficulties that could delay or significantly alter the original agency approved project scope. It should also be noted that at the time of this proposal review, the CES Proposal Review Board apparently did not have members versed in transportation issues or analysis. This translated to a substantial amount effort invested to demonstrate and document Census benefits so that CES Board members were comfortable with the proposed study.

Overall, the CES proposal process itself is somewhat confusing. As noted earlier, during the proposal process, the requisite databases must be identified, but ACS databases are not included in the main pull down list. The ACS data are requested as additional databases, but this field seems to be dropped when the proposal is electronically submitted (or was in the case of this study). As a solution, first, the RDC Director indicated that one of the economic databases had to be specified in the “required” field. This was later amended to preferring that no selection of the economic database occur, only identification of the ACS demographic database. However, the pre-proposal can’t be submitted without selecting an economic database and once a required economic database has been selected there is no way to “unselect” it. The end result in this study was that the RDC Director had to override the web interface and manually allow download and acceptance of the pre-proposal. Recall that the full proposal cannot be submitted until the pre-proposal has been accepted.

³ Holly, B. (2003), *Email Correspondence*, May 16.

This type of confusion highlights the issue associated with the fact that the RDC's are operationalized through the CES. The RDCs are thus mostly geared toward facilitating research with the economic data. With data such as ACS, the RDCs are in the awkward position of having to support research efforts that are not really part of their "job description." Partially to offset this potential problem, the RDC Director made it clear during the proposal process that Census ACS had to provide the necessary staff support and the data products required for analysis, in writing. While this may alleviate some potential problems, others, such as such as the types of results RDCs are allowed to disclose will remain problematic.

In terms of the time required to go through the proposal process, it took nearly 5-6 months to complete all of the required steps and an additional two years to complete all the other related steps (e.g., special sworn status). Note that several weeks of vacation by both the researcher and RDC Director are also captured during this time period. Nonetheless, the process is too long to reasonably complete even a simple study using ACS data at an RDC within a one year timeframe.

Data Management and Analysis Issues

In early 2003, the Census transitioned from a stand alone PC environment to a centralized data access system with thin-client access⁴. The centralized server is UNIX-based, however, this particular study was originally initiated in 2002 when the RDC was functioning under a PC environment and consequently, assumed that certain PC-based GIS products (e.g., a GIS street reference file, StreetMapUSA) were available⁵. No mention was made of the possibility of a thin-client conversation during the proposal review process. The issue is important because projects that started under the PC environment were allowed to finish under the PC. As a result of the long approval process, FHWA (Elaine Murakami) was told that the project must go forward under the UNIX thin client and no exceptions could be made, and then FHWA was subsequently informed that CES/RDC had no budget for the necessary UNIX-based GIS software. The issue was finally resolved when it became apparent, as a result of

⁴ Murakami, E., E. Christopher (2003), *DOT/FHWA's Experience in Attempting to Use the Census Bureau Research Data Centers to Conduct Demographic Analysis with the American Community Survey*, Draft, Provided by E. Murakami

⁵ The street reference file is a TIGER-based dataset extension that operates under the Windows version of ArcView, and converts edge files to compatible shape files (Holly, B. (2003), *Email Correspondence*, May 16)

additional inquiries to an ESRI representative by FHWA, that the Census had an “enterprise-wide” software license covering installation and maintenance of most of the necessary GIS software.

However, availability of the GIS software did not resolve the issue related to the availability of an appropriate street reference file (which was available under the PC environment, but not under UNIX). A limited solution to the UNIX PC incompatibility, creating a street shape file with names but no block numbering, was eventually found with the help of ESRI and FHWA. However, this solution, in turn, gave rise to another significant problem.

The new shape files contained street addresses for all of California, which resulted in a file that was simply too large to access and manipulate under the thin client. The combination of extremely slow speeds and a huge file made working with the converted California shape file impossible. Scripts were then developed to subset only those zipcodes contained within San Francisco County. These scripts were executed and the program repeatedly crashed due to a lack of file space and other problems related to using the converted shapefile. With this solution largely unworkable, a second set of scripts were developed, this time dividing the file into subsets and attempting to work with each subset individually. Again, the program crashed repeatedly. At this point, the testing of workplace address geocoding was dropped as an FHWA research objective. It should be noted that in both cases, scripts were tested successfully on TIGER files on a (fast) UC Davis work station with abundant file storage space. The main problem with the thin client server seems to be related to a combination of very slow computational speeds⁶ and limited storage space. It appeared that each time the program crashed, extraneous files accumulated, further diminishing storage space, and could only be deleted by east coast Census staff.

Another difficulty in data management/analysis related to differing output expectations between the RDC and the Census groups like ACS. The RDC is required to only allow disclosure of statistical model output (e.g., regression or discrete choice models). However, in

⁶ The use of the term “computational speed” refers to both the data transfer speeds and the actual computational speed.

this particular study, it was the availability of robust information⁷ at the appropriate spatial scale that was of interest. That is, the primary focus in this study was to determine if the types of data and numbers of samples collected in each wave of the ACS would constitute a robust enough sample for transportation modeling purposes.

In the case of the disclosure for this project, the conflict between RDC regulations and ACS data analysis needs was highlighted. The RDC expected model output to be submitted for disclosure; this study was looking at, among other things, sample sizes by year, a tabular format. It was difficult to ascertain even the minimum acceptable sample sizes for table cells and a decision was made to disclose results to Census. The disclosure review indicated that the minimum sample sizes in each cell had not been achieved and referred to an internal Census guidance memo on acceptable minimum sample sizes. The RDC Director was unaware of this guidance memo; however, it wouldn't have mattered because the guidance memo wasn't approved for release.

In general, the RDC staff are not very knowledgeable about ACS data (and perhaps other types of non-economic data as well). This might be expected given that the primary mission of the RDCs seems to be oriented toward analysis support of CES data. The availability of in-house resources for addressing questions, such as those related to the acceptable types of ACS output or data weighting, meant that questions could not be readily answered directly in the RDC. Many questions had to be referred to DC and ultimately output was disclosed just so that feedback on minimum sample sizes by Census officials could be provided.

With respect to data and account management, the procedures under the thin client are fairly cumbersome. Passwords must change every month and if there is a problem, the RDC Director can rarely fix the problem directly. For example, during the orientation session, we were unable to establish a password; the problem was related to something that had to be addressed in DC, not Berkeley.

⁷ Robust here refers to minimizing gross errors, rounding and grouping errors, and departure from an assumed sample distribution that may occur with small or biased samples.

Paperwork Issues

Although a certain amount of paperwork is to be expected, the amount required to access ACS data through the RDC's and to pass disclosure is significant. Researchers are required to submit multiple passes of reports and work contracts, all expected to be slightly different in form and content. In addition to these are the normal Special Sworn Status and IRB review applications. It is very clear that CES has instituted a number of repetitive steps designed to expedite studies that deal with the economic data. However, these procedures do not always coincide with the types of expectations that might be associated with analysis of ACS data. For example, census benefits must be enumerated 3-4 times during the proposal process and then again for disclosure. As Table 3 shows, the proportion of project-related time, above that normally associated with project management, was fairly substantial. This almost certainly translates to increased project costs or decreases in project scope.

General Observations

- There is not really any documentation that fully identifies each step in the research process that must be accomplished, and in what order, to satisfy both the RDC requirements and those associated with using the ACS. Many of the forms that must be completed are only released when the prior step has been completed.
- The RDC Director is extremely reliant on the Washington CES office for computer-related administration. Many delays were encountered as a result of lack of coordination between Washington and the RDC. For example, an approximately two-week delay in submitting the pre-proposal occurred apparently because the RDC Director lacked administrator access privileges to the CES website
- It is not clear that the CES Proposal Review Board, as it is currently constituted, is the optimal group for reviewing ACS-related transportation studies. For example, reviewers found the technical (transportation) language very confusing and the value to the Census unclear. While it is understandable, even commendable, that the Census seeks agency-

related benefits from all analyses, this should not preclude recognition of benefits that accrue to other public agencies as sufficient for approval to conduct RDC-based research.

- With a primary focus oriented toward support of CES-related data, there is very little direct staff support available for ACS research. This was a strong concern expressed by the RDC Director when the project was initially approved.
- The RDC computer facility at UC Berkeley is very inadequate. There is no writing space in the facility itself, which hampers onsite documentation of data analysis. IT security upgrades were required before the project could begin and the thin client server is extremely slow and cumbersome.
- Finally, for datasets that fall outside the normal range of RDC experience, such as the ACS, there is a decided lack of coordination between the RDC and CES. One example of this was that only population and household data were delivered to the RDC, despite the explicit request for workplace information. The RDC Director notified CES at least three times prior to the orientation session that the workplace data were not available. Despite this notification, the employment data was not delivered to the RDC until several weeks later.

The use of the RDC for analyzing ACS data is critical. There are questions related to using the ACS data that very important for a range policy issues and exercises, including those associated with transportation planning and modeling and air quality. However, in order to fully utilize the RDC capabilities, some improvements are critical.

The Census must improve the general accessibility of access to microdata for demographic surveys in order to make such studies affordable. This includes clarifying and improving the efficiency of the proposal process, providing more efficient ways in which questions can be resolved during an on-going study, and generally enhancing the work environment at the RDCs (e.g., improving computer technology/server access, workspace, etc.). In particular, the Census should monitor the performance of the thin-client server and make changes as

appropriate. The issue associated with improving access to, and affordability of microdata is not new and is also discussed in Duncan et al.⁸ (1993) and more recently in Schofer et al. (2003)⁹.

⁸ Duncan, G.T., T.B. Jabine, and V.A. de Wolf (1993) *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics*, National Research Council, Washington DC.

⁹ Schofer et al (2003) *Measuring Personal Travel and Goods Movements, A Review of the Bureau of Transportation Statistics Surveys*, National research Council, Special Report 277, Washington, DC.

Appendix A: Detailed Timeline

NOTE: RM – Rich Milby
EM – Elaine Murakami
DN – Deb Niemeier

Project Begin

2/08/02 - FHWA contacts DN regarding the research
2/12/02 - Sent email to Phil requesting clarification on information about ACS
2/19/02 - Clara from Census responded to questions for Phil
2/20/02 - DN sends rough draft of proposal to EM
2/22/02 - EM responded that the proposal looked fine
2/27/02 - Conf call with RM set up

3/06/02 - Conf. call with RM, Elaine, and DN
3/07/02 - Proposal sent to RM for review
3/07/02 - EM requested Phil to send ArcView to RM
3/08/02 - RM send back comments to DN
3/14/02 - DN creates a pre-proposal
3/22/02 - RM told that pre-proposal has been marked so he could accept it
3/25/02 - RM asks DN to make some small changes in proposal
3/26/02 - DN tells RM the changes are done
3/26/02 - RM says to send full proposal to him and to upload files also
3/27/02 - DN sends full proposal to RM for review and uploads file

4/10/02 - RM sends back comments on full proposal
4/15/02 - DN makes changes and uploads new version

5/03/02 - RM asks DN for reviewers for proposal
5/03/02 - DN sends reviewer names

6/17/02 - RM sends reviewer comments to DN, asks for input to take to the meeting
6/18/02 - DN provides responses and modifications

7/03/02 - RM provides formal feedback from review meeting
7/05/02 - EM asks Phil for clarification on review findings
7/08/02 - DN sends revised proposal to Phil, Elaine requests comments
7/16/02 - Everything revised and resubmitted to RM
7/22/02 - DN requests update on proposal status from RDC

8/02/02 - DN requests update on proposal status from RDC
8/02/02 - RM notifies EM that for data that is not part of the CES "core" data, the researcher must take the lead in acquiring it for use and CES must have disclosure agreement in writing.

8/07/04 - Information for submitting for sworn status received from RDC Director
8/23/02 - Special Sworn Status information received
8/30/02 - UCB Human Subjects information received/return

8/30/02 - RDC requests funding payment be submitted

9/03/02 - Elaine requests help from Census ACS for funding transfer to CES

9/11/02 - ACS Census agrees to provide staff support and starts data prep

9/12/02 - DN requests information for livescan fingerprinting

9/19/02 - RM requests livescan information from census

10/05/02 - SSS forms sent by jitney to UCB

10/22/02 - RM emails DN, SSS forms not received

11/12/02 - DN resends all forms via Fed-Ex

11/25/02 - Census ACS delivers ACS data to CES

12/05/02 - RM sends SSS to Wash, DC

1/06/03 - Census requests additional information for special sworn status

2/03/03 - DN completes writing ArcInfo scripts for GIS workplace tasks

2/20/03 - EM requests extension to contract through Oct 03

3/13/03 - Received notification that clearance to work had been approved

3/13/03 - Work contract received from RM, must be completed prior to starting process

3/14/03 - Received confirmation that \$6k would be paid by Census ACS

3/17/03 - Received notification that RDC was now using the Thin Client Server;
ArcView won't work

3/17/03 - Census suggests setting up ArcView on UNIX in Bowie

3/23/03 - Work agreement revised to reflect Census ACS payment of \$6k to CES

4/01/03 - DN completes work agreement and returns to RM including revision for \$6k

4/01/03 - RM notifies DN that there is additional paperwork for Thin Client use

4/15/03 - DN requests update from RM re: thin client paperwork and data availability

4/15/03 - RM notifies DN that thin client server has been installed, but ArcView won't work
CES also still needs to create account

4/16/03 - CES decides that the work agreement that signed on April 1st had to be rewritten
since thin-client was installed shortly before that date.

4/17/03 - CES computer account and data request form received

4/17/03 - RM send CES Researcher Handbook, which must be strictly adhered to

4/29/03 - RM notifies DN that CES has ACS data, ArcView is working but printer is not

5/02/03 - RM notifies DN that StreetmapUSA is not working with UNIX Arcview

5/03/03 - DN notifies RM that street file is necessary

5/12/03 - EM notified by CES that they aren't really sure what is being done and request further
clarification on the use of GIS. A detailed timeline is requested.

5/12/03 - EM responds with reasons why GIS reference file is needed

5/13/03 - RM indicates that he thought Census ACS was supposed take care of GIS needs.

Notes that ACS data have been received. Also states that CES will not proceed without detailed summary of variables to be used and expected matching procedures and timeline.

5/13/03 - Conference call set up between CB and FHWA

6/01/03 - RM/Brian Holly indicate to DN that a pre-visit to UCB is necessary prior to officially starting the project

6/2/03 - Brian Holly notifies RDC that TIGER file has been converted to shape files and is the only way it will work. Now available for use.

6/18/03 - DN visits RDC for required pre-visit. Notes that only household and population record are available for use. There were no workplace records.

6/18/03 - RM notifies Census that workplace are missing.

6/19/03 - EM notifies Census that the file is missing and necessary

6/23/03 - Census notifies EM that geographic coding has been completed and can be picked up Census ACS

6/25/03 - RM notifies DN that workplace data are now available

6/27/03 - DN asks RM for data dictionary

6/27/03 - RM says variable list are confidential and RDC cannot send them out

7/01/03 - Workplace data dictionary received from Census; RDC visit is scheduled for Jul 23

7/18/03 - DN completes all Arcview query scripts

7/18/03 - UCB Human Subjects requests additional information

7/21/03 - RM notifies DN that log in is not working
on thin client

7/25/03 - RDC visit scheduled for Aug 6 and 7.

7/31/03 - RM notifies DN that account problems have been fixed

8/05/03 - RM notifies DN that account is working, but printer is still not functioning

8/06-07/03 - RDC visit - the street files will not work, no way to match using names or zips

9/07/03 - DN notes that thin client is still having problems and DN can't log on. RM has requested assistance from DC.

10/06/03 - SAS not working in the RDC (may have been incorrect SAS call by DN)

1/30/04 - Last run on data completed. RM unable to answer questions regarding sample size restrictions. Decision is made to send everything forward to disclosure.

2/24/04 - Files sent to Census CES/ACS for disclosure

4/20/04- RM notifies DN that disclosure can't be completed - concerns over what the output means, size of cells, and no weighting.

4/22/04 - Project terminated

Appendix B: Disclosure Forms

**Request for Clearance of Research Output
Center for Economic Studies and Research Data Centers**

To be entered by CES:

Project: EXAMINING WORKPLACE GEO-CODING FOR TRAVEL FORECASTING USING THE AMERICAN COMMUNITY SURVEY (#267)
Clearance Request Number: 1
Date: March 9, 2004 (Verbal request made Jan 29, 2004)
Submitted by: Deb Niemeier
Submitted to: Ritch Milby, CCRDC-Berkeley
Cleared for release:
Cleared by: Phil Salopek (& Ritch Milby)

This information supports a request to clear research output for removal from a secure RDC site. It is a companion to the overall project clearance record file and in places it will refer to information contained there.

1. General Information:

- a. Name of this request's subdirectory under the project's main clearance directory (e.g., aug99):
RDC2:/ rdcprojects/cb/cb00267/disclosure
- b. Please describe the outputs you propose to remove. (Use as much space as you need.)

The explanations here should be similar to a table in a journal article.

- c. Please state how the outputs are part of the research project as approved. (Note: If these outputs are described in your proposal, merely refer us there.)

As above, the explanations here should be similar to a table in a journal article.

- d. Use and Presentation of Output: Please indicate how you expect the output to be presented. Check all that apply. *Please enter citations of papers, chapters, dissertations, reports, or presentations into the table in the overall project clearance memo.*

- Journal paper
- Working paper (Don't forget about the CES Discussion Paper series!)
- Dissertation
- Book chapter
- Presentation(s) at conference(s)
- Report (e.g., put out by policy organization)
- Memo for internal use
- Supporting or intermediate output not to be published in any of the above
- Other (specify):

2. Research Output Files: For each *research output file* to be removed, enter the following information

- a. File Name (e.g., output.lst)
- b. Description of file – e.g., tables relating to **household characteristics** or models of ____
- c. Program that produced the file (e.g., output.sas): **SAS**
- d. Number of research sample underlying the the file (from overall project clearance documentation memo):

- e. Number of of corresponding *disclosure analysis file(s)* (e.g., output_disc.lst). This appears on the next page. (Note: if the disclosure information – e.g., firm counts -- appears in this file, it will still appear below; please enter the number.)
- f. Result (we enter this) –Whether file was cleared or not – yes or no.
- g. Comments – any relevant information you or the person who clears the file may wish to add. Note, if a table was not cleared, a replacement table may have been cleared; this will be indicated.

Table 1: Research Output Files

File Number	File Name	Description	Program	Research Sample Number	Disclosure Analysis File Number	Cleared	Comments
1							
2							
3							
4							
5							
6							
7							
8							
9							

3. Disclosure output file(s). Please enter the following information for each file:

- a. File name.
- b. Description of file
- c. Name of program that produced the file. (e.g., output_disc.sas). Note, if the disclosure information is produced in the same program as the research output to be removed, please indicate this.
- d. Number of research sample underlying the file (from overall project clearance documentation memo).
- e. Comment – e.g., if you produced the disclosure output as part of an earlier request to clear output, indicate so here.

Table 2: Disclosure Output Files

File Number	File Name	Description	Program	Research Sample Number	Comments
1					
2					
3					
4					
5					
6					
7					
8					
9					

4. Other information or comments: Please enter any further information you feel is relevant.

Here you should list all the variables that are part of the analysis...providing definitions and any special 'limiting'