Appendix M
Allocation of Missing Place of Work Data in Decennial Censuses and CTPP 2000

Coding place of work locations
The primary process used to assign workers to census geographic areas is called place of work coding by the Census Bureau, although more generally it is called geocoding. In Census processing the work location is based upon both the workplace address and employer name responses from the long-form questionnaires. In the census, geocoding is a two-phase operation. The first phase is an automated or computer-match operation. Records not resolved during this phase move on to the computer-assisted clerical geocoding operation. However, not all workers can be completely coded to all geographic levels. For a variety of reasons some complete answers do not fully match to the coding resource materials, sometimes the answers provided are incomplete, and other times no answers are provided at all. So at the end of the geocoding process, some work location geocodes remain missing.

Place of work allocation
In decennial census terms, allocation is the process used to assign data attributes wherever the responses to census questions are either incomplete or missing. The substituted responses are taken from other people who responded to the given question, and were similar in characteristics and residential location to the person with missing information.

Place of work allocation procedures employed by the Census Bureau use both trip data and job data to assign workplace locations. The general process is to find workers whose workplace is fully coded ("Donors") and are similar in characteristics to the workers needing place of work allocation ("Recipients"). If a full match cannot be achieved the rules are relaxed, one by one, in an attempt to find a reasonable donor. However, in some cases, even with relaxed requirements, a donor cannot be found, resulting in missing work location geocodes.

The allocation process is conducted in two steps (1) standard allocation and (2) extended allocation. Standard allocation assigns work locations, at a minimum, to a state, county and place geocode. Many records are also allocated down to the tract, block group and TAZ level during the standard allocation. The extended allocation system goes beyond the standard process to assign even more work locations to tract, block group and TAZ.

Standard allocation in Census 2000
When allocating state and county of work codes, the standard allocation system attempts to match four worker characteristics at once:
- travel time (less than 30 minutes/30 mins. and over)
- residence tract
- means of transportation (public/non-public)
- industry (20 classes).

For each recipient, the procedure begins by looking for a donor that matches the recipient on all four characteristics. If a donor cannot be found, the matching rules are relaxed, one by one, by looking for donors in the other categories of each characteristic. First, the travel time match is relaxed, then the residence tract match, then means of transportation, and finally the industry match is relaxed.

The standard allocation system uses a slightly different process to assign place, tract, and block geocodes to workers. Instead of looking for donors who <u>live</u> near the recipient, the match finds donors who <u>work</u> in the same state and county and applies their place, tract, and block of work codes to the recipients. For these detailed levels of workplace allocation only two match characteristics are used, means of transportation and industry.

In Census 2000, a match is not forced using the standard allocation process. If a match has not been achieved after the last characteristic has been tried by itself, the process is stopped. Thus, while the output from the standard allocation system insures that every worker has workplace codes for at least state, county, and place geography, there is no assurance that tract and block codes will be present.

Extended allocation in CTPP 2000
The extended allocation procedure was developed for CTPP 2000 to reduce the number of missing workplace geocodes for tracts and blocks that remained after the standard allocation process. If State or County codes were assigned in the standard allocation process, those codes are not changed in the extended allocation process. On occasion, a workplace MCD or place code may be changed in extended allocation.

Extended allocation is done in two stages. In the first stage, a set of potential destination areas is demarcated for each recipient, based on trip characteristics and residence location. In the second stage, the recipient is matched to a fully geocoded donor who matches the recipient's industry and occupation characteristics and who works in any one of the potential destination areas.

The first stage begins with the residence tract of the recipient. All residents of that tract who have been coded to a place of work tract are identified, and their work tracts are listed. In addition, geocoded workers who <u>work</u> in that residence tract are identified. Their origin tract numbers are then added to the list of potential work tracts described above. The assumption here is that if workers are able to travel from tract B to tract A, it is also reasonable to expect that workers could travel from tract A to tract B.

Extended allocation uses ten means (mode) of transportation categories and seven travel time (duration) classes. This differs from the standard allocation, which uses only 2 categories of means of transportation and 2 time (duration) classes. The worker being allocated is matched to geocoded workers who live in the same residence tract and who have the same mode and travel time. The work tracts of these coded workers provide the set of destination areas for the worker being allocated. If an exact match on mode and duration cannot be made, the rules are relaxed.

The rules are relaxed through the use of substitute mode and travel time classes. For eight of the ten original modes (ferryboat and other means are excluded), there is a primary substitute mode. For example, the primary substitute mode for taxicab is car/truck/van/motorcycle, and vice versa. In addition, the four transit modes (streetcar, bus, railroad, and subway) have a secondary substitute mode, called public transit. Travel time also has substitute classes, called contiguous durations. If a match cannot be achieved using the given travel time, for example 30-44 minutes, matching will be attempted using travel times of 20-29 minutes and 45-59 minutes.

In the second stage, industry and occupation are used to select a specific donor (and tract/block) from the collection of workers who worked in the potential destination tracts. Industrial classes used in the match are in a three-level hierarchy, beginning with a detailed breakdown of 90 classes. The standard allocation uses only 20 industry classes. Occupational classes are in a two-level hierarchy, beginning with a detailed breakdown of 24 classes. If more than one donor match on industry and occupation is found from among the coded workers working in the potential destination tracts, one is randomly chosen and the process is complete for the recipient.